

**KOMPARASI METODE *INVERSE-TRANSFORM*, *ACCEPTANCE-REJECTION*, DAN *TRANSFORMATION (BOX-MULLER)* DALAM  
SIMULASI DATA PENUMPANG ANGKUTAN UDARA SUMATERA  
SELATAN**

TUGAS BESAR MATA KULIAH KOMPUTASI STATISTIK



Kelompok 8

Disusun Oleh:

Giofani Aristyo                      123450065

Ahmad Rizky                        123450050

Romauli Oktavia Silaban    123450014

Tesalonika Hutajulu            123450033

**PROGRAM STUDI SAINS DATA  
FAKULTAS SAINS  
INSTITUT TEKNOLOGI SUMATERA  
LAMPUNG SELATAN**

**2025**

## DAFTAR PUSTAKA

BAB I PENDAHULUAN.....	5
1.1 Latar Belakang .....	5
1.2 Rumusan Masalah .....	5
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	6
BAB II TINJAUAN PUSTAKA.....	7
2.1    Landasan Teori Distribusi Peluang dalam Simulasi .....	7
2.1.1. Distribusi Normal .....	7
2.1.2. Distribusi Log-Normal .....	7
2.1.3. Distribusi Gamma.....	8
2.2    Estimasi Parameter dengan <i>Maximum Likelihood Estimation</i> .....	9
2.3    Pembangkit Bilangan Acak Dasar Pada Bahasa Pemograman R .....	10
2.4    Metode Pembangkit Bilangan Acak.....	10
2.4.1. <i>Inverse-Transform</i> .....	11
2.4.2. <i>Acceptance-Rejection</i> dengan <i>Cauchy Proposal</i> .....	11
2.4.3. <i>Transformation Method</i> dengan <i>Box-Muller</i> .....	11
2.5    Uji Validasi Model ( <i>Goodness of Fit</i> ).....	12
2.5.1. Uji Kolmogorov-Smirnov (KS-Test).....	12
2.5.2. Evaluasi Momen Statistik.....	12
BAB III METODE PENELITIAN .....	14
3.1.    Jenis dan Sumber Data.....	14
3.2.    Teknik Pengumpulan Data .....	14
3.3.    Variabel Penelitian .....	14
3.4.    Tahapan Penelitian .....	17
BAB IV HASIL DAN PEMBAHASAN .....	19
4.1.    Eksplorasi Data Penumpang Angkutan Udara.....	19
4.1.1. Pola Pergerakan dan Dekomposisi Deret Waktu .....	19
4.1.2. Karakteristik Distribusi Data .....	20
4.2.    Uji Asumsi Distribusi dan Estimasi Parameter.....	22
4.2.1. Uji Normalitas Data.....	22
4.2.2. Estimasi Parameter ( <i>Maximum Likelihood Estimation</i> ) .....	23
4.3.    Implementasi Simulasi Pembangkit Bilangan Acak .....	23
4.3.1    Metode <i>Inverse-Transform</i> .....	23

4.3.2	Metode <i>Acceptance-Rejection</i> .....	27
4.3.3	<i>Transformation Method (Box-Muller)</i> .....	30
4.4.	Evaluasi dan Perbandingan Kinerja Metode.....	34
4.4.1.	Analisis Bias Komparatif.....	35
4.4.2.	Tinjauan Goodness-of-Fit .....	36
4.4.3.	Pemilihan Metode Terbaik.....	36
BAB V KESIMPULAN.....		37
5.1.	Kesimpulan .....	37
5.2.	Saran .....	37
DAFTAR PUSTAKA.....		38

## ABSTRAK

Transportasi udara di Sumatera Selatan memegang peranan vital dalam mobilitas ekonomi namun memiliki karakteristik data penumpang yang sangat fluktuatif akibat pengaruh musiman dan guncangan eksternal pandemi COVID-19. Ketidakpastian ini menuntut pendekatan simulasi stokastik yang akurat untuk mendukung perencanaan operasional bandara. Penelitian ini bertujuan untuk membandingkan kinerja tiga metode pembangkitan bilangan acak, yaitu *Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation Method* (Box-Muller), dalam mensimulasikan data penumpang angkutan udara periode 2015–2025. Data sekunder diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan dan teridentifikasi mengikuti distribusi Normal berdasarkan uji Shapiro-Wilk. Simulasi dilakukan dengan memvariasikan ukuran sampel ( $N = 100, 1.000, 10.000$ ) untuk menguji konsistensi estimator. Evaluasi kinerja metode didasarkan pada uji validasi distribusi *Kolmogorov-Smirnov* (KS-Test) dan analisis Bias parameter (*Mean* dan *Standard Deviation*). Hasil penelitian menunjukkan bahwa ketiga metode valid secara statistik dalam mereplikasi distribusi data asli ( $p\text{-value} > 0.05$ ). Metode *Acceptance-Rejection* menghasilkan kesesuaian bentuk distribusi terbaik dengan  $p\text{-value}$  tertinggi (0,7121). Namun, metode *Box-Muller* terbukti menjadi metode yang paling presisi dan *robust* dengan bias standar deviasi terendah, mencapai angka 38,98 pada sampel besar ( $N = 10.000$ ), jauh lebih kecil dibandingkan metode lainnya. Oleh karena itu, metode *Box-Muller* direkomendasikan sebagai pendekatan terbaik untuk simulasi data penumpang angkutan udara di Sumatera Selatan.

Kata Kunci: Simulasi, *Inverse-Transform*, *Acceptance-Rejection*, *Box-Muller*, Angkutan Udara, BPS Sumatera Selatan.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Transportasi udara merupakan salah satu moda yang memiliki peranan krusial dalam mendukung konektivitas dan mobilitas masyarakat, serta menunjang kegiatan ekonomi di wilayah Sumatera Selatan. Sektor ini menjadi katalisator pertumbuhan ekonomi wilayah melalui percepatan arus barang dan jasa [1]. Namun, berdasarkan laporan tahunan Badan Pusat Statistik, jumlah penumpang angkutan udara yang datang dan berangkat selama periode 2015 hingga 2025 menunjukkan pola yang sangat fluktuatif [2]. Variasi ini dipengaruhi oleh faktor musiman (*seasonality*) seperti libur hari raya, kondisi makroekonomi, hingga guncangan eksternal (*shock*) seperti pandemi COVID-19 yang menyebabkan perubahan struktural pada data runtun waktu [2]. Ketidakstabilan ini menuntut analisis yang mendalam karena pendekatan deterministik sering kali gagal menangkap dinamika ketidakpastian tersebut.

Kondisi ini memberikan tantangan tersendiri bagi pengelola bandara dan pemangku kebijakan transportasi di Sumatera Selatan dalam menyusun perencanaan operasional, seperti penentuan kapasitas dan penjadwalan. Menurut Horonjeff et al. [3], perencanaan fasilitas bandara yang efisien memerlukan estimasi permintaan yang akurat untuk meminimalisir kongesti maupun *over-capacity*. Oleh karena itu, diperlukan metode statistika komputasi yang mampu memodelkan dan mensimulasikan variasi alami pergerakan penumpang tersebut. Pembangkitan bilangan acak (*random number generation*) menjadi pondasi utama dalam membangun simulasi stokastik untuk meniru perilaku sistem nyata yang mengandung ketidakpastian [4].

Dalam penelitian ini, simulasi dilakukan menggunakan tiga pendekatan algoritma, yaitu *Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation Method* (Box-Muller). Metode *Inverse-Transform* dan *Acceptance-Rejection* merupakan teknik fundamental dalam membangkitkan variabel acak dari distribusi tertentu, sedangkan metode Box-Muller dikenal efisien secara komputasi untuk membangkitkan distribusi Normal standar [6] [7]. Penerapan ketiga metode ini bertujuan untuk merepresentasikan ketidakpastian data penumpang angkutan udara secara realistis. Dengan membandingkan kinerja ketiga metode tersebut, diharapkan dapat diperoleh model simulasi yang paling akurat (*robust*) untuk memberikan gambaran komprehensif mengenai pola pergerakan penumpang, sehingga dapat mendukung pengambilan keputusan yang lebih berbasis data.

### 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, permasalahan utama yang akan dikaji dalam penelitian ini dapat dirumuskan ke dalam beberapa pertanyaan penelitian sebagai berikut:

1. Bagaimanakah karakteristik distribusi dan parameter data jumlah penumpang angkutan udara di Sumatera Selatan periode 2015–2025?

2. Bagaimanakah implementasi algoritma pembangkitan bilangan acak menggunakan metode *Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation (Box-Muller)* untuk mensimulasikan data tersebut?
3. Metode manakah yang memberikan kinerja terbaik dalam membangkitkan data simulasi berdasarkan kriteria akurasi parameter (bias) dan kesesuaian distribusi (goodness of fit)?

### 1.3 Tujuan Penelitian

Mengacu pada rumusan masalah yang telah dipaparkan sebelumnya, penelitian ini dilaksanakan dengan tujuan-tujuan sebagai berikut:

1. Mengidentifikasi pola distribusi, asumsi normalitas, dan estimasi parameter dari data jumlah penumpang angkutan udara.
2. Mengembangkan kode program dalam perangkat lunak R untuk membangkitkan bilangan acak menggunakan metode *Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation (Box-Muller)*.
3. Mengevaluasi dan membandingkan efektivitas ketiga metode simulasi tersebut untuk menentukan metode yang paling presisi dalam mereplikasi karakteristik data penumpang.

### 1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi positif baik dari sisi pengembangan ilmu pengetahuan maupun penerapannya di lapangan, antara lain:

1. Manfaat Teoritis
  - a. Menambah khazanah pengetahuan mengenai penerapan metode pembangkitan bilangan acak (*random number generation*) pada kasus data riil yang memiliki karakteristik kompleks (fluktuatif dan musiman).
  - b. Memberikan bukti empiris mengenai kelebihan dan kekurangan dari metode *Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation (Box-Muller)* dalam hal akurasi parameter dan efisiensi komputasi, sehingga dapat menjadi referensi bagi studi simulasi selanjutnya.
2. Manfaat Praktis
  - a. Hasil simulasi dapat digunakan sebagai bahan pertimbangan dalam perencanaan operasional yang berbasis risiko (*risk-based planning*), seperti estimasi kapasitas penumpang di masa puncak (*peak season*) guna mengoptimalkan penjadwalan petugas dan fasilitas bandara.
  - b. Memberikan wawasan mengenai pola ketidakpastian mobilitas angkutan udara di Sumatera Selatan, yang dapat dimanfaatkan sebagai data pendukung dalam perumusan kebijakan transportasi daerah pasca-pandemi.
  - c. Mengasah kemampuan dalam mengimplementasikan algoritma statistik ke dalam bahasa pemrograman R serta meningkatkan ketajaman analisis dalam menerjemahkan fenomena data menjadi model simulasi yang logis.

## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Landasan Teori Distribusi Peluang dalam Simulasi

Dalam simulasi stokastik sistem transportasi udara, ketidakpastian permintaan merupakan aspek fundamental yang memerlukan pemodelan probabilistik yang cermat [8]. Sistem transportasi dicirikan oleh sejumlah variabel probabilistik, seperti frekuensi kedatangan penumpang dan variasi waktu layanan, yang tidak dapat diprediksi secara pasti melalui pendekatan deterministik. Proses pemodelan yang akurat mensyaratkan identifikasi dan penerapan distribusi probabilitas teoritis yang sesuai, yang mampu menghasilkan bilangan acak yang mencerminkan pola dan karakteristik empiris dari data aktual. Pendekatan ini memungkinkan peneliti dan perencana sistem transportasi untuk:

1. Mengembangkan simulasi yang lebih realistis
2. Menangkap variabilitas alami dalam sistem transportasi
3. Menghasilkan proyeksi yang lebih andal tentang kinerja dan kapasitas

Inti dari metode ini terletak pada kemampuan untuk mentransformasi data empiris menjadi model probabilistik yang dapat mensimulasikan berbagai skenario potensial dalam sistem transportasi udara.

##### 2.1.1. Distribusi Normal

Distribusi Normal merupakan instrumen statistik yang paling fundamental, disahkan oleh Teorema Limit Pusat yang menjelaskan konvergensi rata-rata variabel acak independen menuju distribusi normal, tanpa mempertimbangkan distribusi aslinya. Dalam ranah analisis transportasi udara, distribusi ini berperan sebagai alat pemodelan strategis untuk mengestimasi agregasi penumpang selama interval waktu yang panjang, khususnya pada kondisi dengan probabilitas rendah terhadap fluktuasi signifikan. Karakteristik kritis distribusi normal terletak pada kemampuannya menghasilkan nilai dalam rentang tak terhingga ( $-\infty$  sampai  $+\infty$ ), yang secara paradoks juga menjadi kelemahannya dalam konteks pemodelan fenomena dengan batasan non-negatif. Untuk menjembatani kesenjangan ini, para analis statistik lazimnya menggunakan teknik pemotongan (*truncation*), membatasi distribusi pada nilai di atas nol guna memastikan representasi matematis yang lebih akurat untuk skenario seperti perhitungan jumlah penumpang. Fungsi kepadatan peluang (PDF) untuk distribusi normal didefinisikan sebagai:

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

Di mana  $\mu$  adalah rata-rata (*mean*) dan  $\sigma$  adalah standar deviasi [9].

##### 2.1.2. Distribusi Log-Normal

Analisis empiris dalam domain transportasi mengungkapkan pola kompleks dari data permintaan perjalanan, yang kerap menampilkan karakteristik kemencengan positif (*right-skewed*). Distribusi Log-Normal emerges sebagai instrumen statistik yang sangat tepat untuk memetakan variabel dengan batasan non-negatif dan volatilitas tinggi, seperti dinamika

pendapatan penumpang atau volume lalu lintas udara di kawasan dengan pertumbuhan cepat. Secara matematis, distribusi Log-Normal memiliki hubungan fundamental dengan distribusi normal: ketika variabel acak  $Y$  mengikuti distribusi normal, maka transformasi eksponensial dari  $Y$  akan menghasilkan distribusi Log-Normal. Dalam konteks spesifik wilayah berkembang seperti Sumatera Bagian Selatan, distribusi ini memperlihatkan keunggulan signifikan dibandingkan distribusi normal. Keistimewaan distribusi Log-Normal terletak pada kemampuannya untuk:

1. Mengakomodasi lonjakan ekstrem dalam jumlah penumpang pada musim-musim tertentu.
2. Menghindari munculnya nilai-nilai negatif dalam simulasi.
3. Merepresentasikan variabilitas tinggi pada sistem transportasi yang dinamis

Fungsi kepadatan peluangnya adalah:

$$f(x) = \frac{1}{x \sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0$$

Di mana  $\mu$  merepresentasikan parameter rata-rata dari logaritma variabel dan  $\sigma$  merepresentasikan standar deviasi dari logaritma variable.

### 2.1.3. Distribusi Gamma

Distribusi Gamma muncul sebagai instrumen statistik yang sangat adaptif dalam domain teori antrian, dengan fokus utama pada pemodelan probabilistik waktu tunggu dan interval antar kedatangan penumpang [10]. Keunggulan fundamental distribusi ini terletak pada fleksibilitas strukturalnya, yang ditentukan oleh dua parameter kunci: parameter bentuk ( $\alpha$ ) dan parameter skala ( $\beta / \theta$ ). Dalam konteks simulasi operasional bandara, distribusi Gamma menunjukkan kapasitas komputasional yang luar biasa. Kemampuannya untuk merepresentasikan spektrum kurva probabilitas membuat distribusi ini menjadi pilihan ideal untuk menangkap kompleksitas pola kedatangan penumpang yang tidak bersifat seragam. Karakteristik unik distribusi Gamma mencakup:

1. Adaptabilitas morfologis dari kurva probabilitas.
2. Transisi dari bentuk eksponensial ( $\alpha = 1$ ) hingga mendekati distribusi normal (untuk  $\alpha$  yang besar)
3. Representasi akurat terhadap variabilitas temporal dalam sistem transportasi

Dengan demikian, distribusi Gamma menawarkan kerangka matematis yang komprehensif untuk menganalisis dan mensimulasikan dinamika kompleks dalam sistem transportasi udara, khususnya pada skenario dengan pola kedatangan yang tidak linier. PDF distribusi gamma dinyatakan sebagai:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0$$

Dalam persamaan fungsi kepadatan peluang di atas, variabel  $x$  merepresentasikan variabel acak kontinu yang bernilai positif ( $x > 0$ ). Simbol  $\alpha$  (alpha) merupakan parameter bentuk (*shape*



*parameter*) yang menentukan profil kemiringan kurva, sedangkan  $\beta$  (beta) berfungsi sebagai parameter laju (*rate parameter*) yang mengendalikan penyebaran data<sup>1</sup>. Adapun  $\Gamma(\alpha)$  adalah fungsi Gamma, yang merupakan bentuk generalisasi faktorial untuk bilangan riil dan berfungsi sebagai konstanta normalisasi agar total probabilitas di bawah kurva sama dengan satu.

## 2.2 Estimasi Parameter dengan *Maximum Likelihood Estimation*

Setelah identifikasi distribusi probabilitas yang paling representatif untuk data penumpang angkutan udara, proses pemodelan simulasi berlanjut pada tahap krusial: estimasi parameter distribusi terpilih. Baik untuk distribusi normal (dengan parameter  $\mu$  dan  $\sigma$ ) maupun gamma (dengan parameter  $\alpha$  dan  $\beta$ ), metode estimasi standar yang memiliki kredibilitas statistik tinggi adalah *Maximum Likelihood Estimation* (MLE). Esensi MLE terletak pada pendekatan sistematis untuk mengekstraksi parameter model yang paling optimal [9]. Metode ini beroperasi melalui optimasi fungsi *likelihood*, yakni fungsi matematis yang mengukur probabilitas kemunculan sekumpulan data observasi berdasarkan parameter tertentu. Secara filosofis, MLE bertujuan mengidentifikasi kombinasi parameter yang membuat data sampel yang diamati menjadi paling mungkin terjadi. Keunggulan metode MLE mencakup:

1. Kemampuan menghasilkan estimasi parameter yang konsisten.
2. Memiliki sifat statistik asimptotik yang kuat.
3. Fleksibilitas dalam berbagai jenis distribusi probabilitas.
4. Akurasi tinggi dalam mengekstrapolasi karakteristik populasi dari sampel terbatas.

Dengan demikian, MLE menyediakan kerangka komputasional yang canggih untuk mentransformasi data empiris menjadi model probabilistik yang presisi dalam simulasi transportasi udara.

Jika  $X_1, X_2, \dots, X_n$  adalah sampel acak independen dari populasi dengan fungsi kepadatan peluang  $f(x; \theta)$ , maka fungsi *likelihood*  $L(\theta)$  didefinisikan sebagai perkalian dari kepadatan peluang setiap observasi:

$$L(\theta) = L(\theta; x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta)$$

Dalam estimasi parameter statistik, fungsi *likelihood* yang melibatkan perkalian kompleks diselesaikan melalui transformasi logaritmik. *Log-likelihood* mengubah perkalian menjadi penjumlahan, dengan memanfaatkan sifat fungsi logaritma natural sebagai fungsi monoton naik. Dengan demikian, titik maksimum fungsi *likelihood*  $L(\theta)$  akan sama dengan titik maksimum *log-likelihood*  $\ln L(\theta)$ , namun dengan perhitungan yang jauh lebih sederhana dan mudah dikomputasi. Pendekatan ini memungkinkan peneliti untuk dengan efisien mengekstraksi estimasi parameter distribusi probabilitas dalam berbagai model simulasi.

$$l(\theta) = \ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta)$$

Estimator MLE, dilambangkan dengan  $\hat{\theta}$ , diperoleh dengan mencari turunan pertama dari fungsi *log-likelihood* terhadap parameter  $\theta$  dan menyamakannya dengan nol:

$$\frac{\partial l(\theta)}{\partial \theta} = 0$$

Dalam konteks data penumpang angkutan udara:

1. Untuk Distribusi Normal: MLE akan menghasilkan estimator rata-rata sampel ( $\bar{x}$ ) sebagai penaksir  $\mu$ , dan varians sampel sebagai penaksir  $\sigma^2$ .
2. Untuk Distribusi Eksponensial/Gamma: MLE digunakan untuk mencari laju kedatangan ( $\lambda$ ) atau parameter bentuk yang paling sesuai dengan pola kedatangan penumpang di bandara.

## 2.3 Pembangkit Bilangan Acak Dasar Pada Bahasa Pemrograman R

Simulasi statistik pada komputer bersifat deterministik, yang berarti komputer tidak dapat menghasilkan keacakan yang murni. Oleh karena itu, simulasi menggunakan algoritma *Pseudo-Random Number Generator* (PRNG) untuk menghasilkan urutan angka yang secara statistik berperilaku seperti variabel acak, meskipun sebenarnya dihasilkan oleh rumus matematika yang pasti.

Dalam lingkungan komputasi statistik R, algoritma standar (*default*) yang digunakan untuk membangkitkan bilangan acak seragam kontinu  $U(0,1)$  adalah *Mersenne Twister* (MT19937) [11]. Algoritma ini dikembangkan oleh Matsumoto dan Nishimura pada tahun 1997 dan dipilih sebagai standar industri karena tiga keunggulan utamanya [12]:

1. Mersenne Twister memiliki periode  $2^{19937} - 1$ . Periode ini menjamin bahwa dalam simulasi penumpang angkutan udara yang melibatkan ribuan atau jutaan iterasi sekalipun, pola angka tidak akan berulang, sehingga menjaga validitas statistik simulasi.
2. Algoritma ini terbukti memiliki ekuidistribusi hingga 623 dimensi, yang berarti korelasi antar bilangan acak yang dihasilkan sangat kecil (dapat diabaikan).
3. Algoritma ini sangat cepat dan efisien dalam penggunaan memori, sehingga cocok untuk simulasi skala besar.

Proses pembangkitan dimulai dengan nilai awal yang disebut *seed*. Jika  $X_0$  adalah *seed*, maka algoritma PRNG akan menghasilkan barisan  $X_n$  berdasarkan relasi rekursif:

$$X_{n+1} = f(X_n, \theta)$$

Di mana  $f$  adalah fungsi algoritma *Mersenne Twister* dan  $\theta$  adalah parameter algoritma. Nilai output kemudian dinormalisasi menjadi rentang (0, 1) untuk mendapatkan variabel acak seragam  $U$ , yang nantinya akan ditransformasikan menjadi distribusi target (seperti normal atau gamma) yang merepresentasikan data penumpang.

## 2.4 Metode Pembangkit Bilangan Acak

Sesudah bilangan acak seragam  $U(0,1)$  diperoleh dari generator dasar, untuk langkah selanjutnya dalam simulasi adalah mengubah bilangan tersebut menjadi variabel acak yang mengikuti distribusi target (misalnya normal untuk berat bagasi atau gamma untuk waktu

kedatangan penumpang). Terdapat beberapa metode standar yang digunakan dalam simulasi statistik, di antaranya adalah Metode Transformasi Invers, Metode *Box-Muller*, dan metode *Acceptance-Rejection* [12].

#### **2.4.1. Inverse-Transform**

Metode ini didasarkan pada prinsip probabilitas integral transform. Jika  $X$  adalah variabel acak kontinu dengan Fungsi Distribusi Kumulatif (CDF)  $F(X)$  yang dapat dibalik (invertible), maka variabel acak  $U$  dapat didefinisikan sebagai  $U = F(X)$ . Karena  $U$  berdistribusi seragam pada interval  $(0,1)$ , maka kita dapat membangkitkan  $X$  dengan menggunakan invers dari CDF tersebut [5]:

$$X = F^{-1}(U), \quad \text{di mana } U \sim U(0,1)$$

Metode ini sangat efisien untuk distribusi yang memiliki bentuk invers CDF tertutup, seperti distribusi eksponensial. Namun, untuk distribusi seperti normal yang tidak memiliki bentuk invers tertutup sederhana, metode ini memerlukan pendekatan numerik yang lebih kompleks, sehingga sering kali metode lain lebih disukai.

#### **2.4.2. Acceptance-Rejection dengan Cauchy Proposal**

Metode *Acceptance-Rejection* digunakan ketika fungsi kepadatan peluang  $f(x)$  dari distribusi target sulit untuk disimulasikan secara langsung (misalnya invers-nya sulit dihitung). Konsepnya adalah menggunakan distribusi proposal  $g(x)$  yang lebih mudah disimulasikan dan "menyelimuti" distribusi target, sedemikian rupa sehingga  $f(x) \leq c \cdot g(x)$  untuk konstanta  $c \geq 1$  [13]. Salah satu pendekatan umum untuk membangkitkan distribusi normal atau distribusi berekor tebal lainnya adalah menggunakan distribusi Cauchy sebagai proposal. distribusi Cauchy dipilih karena memiliki ekor yang lebih berat (*heavier tails*) dibandingkan distribusi normal, sehingga memenuhi syarat amplop (*envelope*) yang lebih aman untuk menghindari area yang tidak tercover. Algoritmanya adalah:

1. Bangkitkan  $Y$  dari distribusi Cauchy (proposal).
2. Bangkitkan  $U \sim U(0,1)$
3. Jika  $U \leq \frac{f(Y)}{c \cdot g(Y)}$ , maka terima  $Y$  sebagai  $X$ . Jika tidak, tolak dan ulangi.

#### **2.4.3. Transformation Method dengan Box-Muller**

Metode Box-Muller (1958) adalah algoritma spesifik yang sangat efisien untuk membangkitkan pasangan bilangan acak berdistribusi normal standar dari dua bilangan acak seragam independen  $U_1$  dan  $U_2$  [7]. Metode ini memanfaatkan transformasi koordinat dari kartesius ke polar. Rumus pembangkitannya adalah:

$$Z_1 = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$$

$$Z_2 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$$

Dalam konteks simulasi penumpang, nilai  $Z$  ini kemudian dapat disesuaikan (diskalakan) menjadi  $X = \mu + Z\sigma$  untuk merepresentasikan data nyata, seperti variasi jumlah penumpang harian yang mengikuti distribusi normal.

## 2.5 Uji Validasi Model (*Goodness of Fit*)

Validasi model merupakan tahapan krusial dalam studi simulasi untuk memastikan bahwa model yang dibangun merupakan representasi yang akurat dari sistem nyata [14]. Dalam konteks simulasi jumlah penumpang angkutan udara, validasi bertujuan untuk menguji apakah distribusi data hasil pembangkitan bilangan acak memiliki karakteristik statistik yang sama dengan data historis penumpang di lapangan. Proses validasi ini dilakukan melalui uji statistik formal dan evaluasi momen.

### 2.5.1. Uji Kolmogorov-Smirnov (KS-Test)

Uji Kolmogorov-Smirnov (KS-Test) adalah metode non-parametrik yang digunakan untuk menguji kesesuaian antara distribusi empiris sampel (data hasil simulasi) dengan distribusi teoretis tertentu (data target), atau untuk membandingkan dua distribusi sampel [15]. Berbeda dengan uji Chi-Square yang sensitif terhadap pengelompokan data (binning), KS-Test bekerja berdasarkan fungsi distribusi kumulatif (CDF) dan cocok untuk data kontinu. Statistik uji  $D$  didefinisikan sebagai jarak vertikal maksimum antara CDF empiris  $S_N(x)$  dan CDF teoretis  $F(x)$ :

$$D = \sup_x |F_n(x) - F(x)|$$

Hipotesis yang digunakan adalah:

$H_0$ : Data simulasi berdistribusi sesuai dengan adanya data historis.

$H_1$ : Data simulasi tidak mengikuti distribusi adanya data historis.

Jika nilai statistik hitung  $D$  lebih kecil dari nilai kritis tabel pada tingkat signifikansi  $\alpha$  tertentu (biasanya 0.05), maka  $H_0$  diterima, yang berarti data simulasi valid dan representatif.

### 2.5.2. Evaluasi Momen Statistik

Selain uji formal, validasi juga perlu dilakukan dengan membandingkan momen statistik dari data simulasi ( $Y_{sim}$ ) terhadap data riil ( $Y_{real}$ ). Hal ini penting karena KS-Test hanya mengukur jarak maksimum, namun terkadang gagal menangkap detail bentuk distribusi pada bagian ekor (*tail*). Indikator yang dievaluasi meliputi:

1. Rata-rata adalah indikator pemusatan data yang merepresentasikan ekspektasi jumlah penumpang dalam satu periode waktu. Dalam validasi simulasi, selisih antara rata-rata data hasil simulasi ( $\overline{Y_{sim}}$ ) dengan rata-rata data historis ( $\overline{Y_{real}}$ ) disebut sebagai bias. Bias yang bernilai nol atau mendekati nol mengindikasikan bahwa model simulasi memiliki akurasi yang tinggi (*unbiased*). Sebaliknya, bias yang besar menunjukkan adanya kesalahan sistematis, di mana model cenderung melakukan *overestimation* (prediksi terlalu tinggi) atau *underestimation* (prediksi terlalu rendah) terhadap volume penumpang. Rumus evaluasi bias rata-rata dinyatakan sebagai:

$$\text{Bias} = \bar{Y}_{sim} - \bar{Y}_{real}$$

atau sering dinyatakan dalam persentase kesalahan relatif (*Relative Error*):

$$RE_{\mu} = \left| \frac{\bar{Y}_{sim} - \bar{Y}_{real}}{\bar{Y}_{real}} \right| \times 100\%$$

2. Standar deviasi ( $\sigma$ ) mengukur seberapa jauh data tersebar dari rata-ratanya. Dalam konteks angkutan udara di Sumatera Selatan, standar deviasi merepresentasikan fluktuasi atau ketidakpastian jumlah penumpang (misalnya perbedaan drastis antara hari biasa dan musim liburan). Model simulasi yang baik harus mampu menghasilkan standar deviasi ( $S_{sim}$ ) yang mendekati standar deviasi data historis ( $S_{real}$ ). Jika ( $S_{sim}$ ) jauh lebih kecil dari ( $S_{real}$ ), model dianggap "terlalu stabil" dan gagal menangkap dinamika lonjakan penumpang yang sebenarnya terjadi di lapangan. Perbandingan rasio standar deviasi dapat digunakan sebagai indikator validitas variabilitas:

$$\text{Rasio Variabilitas} = \frac{S_{sim}}{S_{real}}$$

Nilai rasio yang mendekati 1 menunjukkan bahwa simulasi berhasil mereplikasi tingkat ketidakpastian sistem nyata dengan baik.

## BAB III

### METODE PENELITIAN

#### 3.1. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder merupakan data yang diperoleh tidak secara langsung dari objek penelitian, melainkan lewat pihak lain atau dokumen yang sudah dipublikasikan. Sumber data dalam penelitian ini berasal dari publikasi resmi Badan Pusat Statistik (BPS) Provinsi Sumatera Selatan. Data yang diambil merupakan data runtun waktu (*time series*) bulanan mengenai Jumlah Penumpang Angkutan Udara yang datang dan berangkat melalui bandara-bandara di wilayah Sumatera Selatan. Periode data yang diamati mencakup rentang waktu 10 tahun, mulai dari tahun 2015 hingga 2025.

#### 3.2. Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan dengan metode dokumentasi dan studi pustaka. Penulis mengakses portal data resmi BPS Sumatera Selatan, mengunduh laporan statistik transportasi, dan melakukan kompilasi data mentah ke dalam format terstruktur (*comma-separated values* atau .csv) agar dapat diproses lebih lanjut menggunakan perangkat lunak statistik.

#### 3.3. Variabel Penelitian

Variabel yang diamati dalam penelitian ini merujuk pada komponen atau karakteristik tertentu yang diukur untuk memperoleh informasi terkait fenomena yang diteliti. Pada konteks penelitian ini, variabel yang diamati mencakup data yang merepresentasikan jumlah penumpang angkutan udara Sumatera Selatan selama tahun 2015-2025. Pengamatan terhadap variabel ini dilakukan untuk memahami dan mengidentifikasi pola distribusi data, serta mendukung proses analisis statistik yang akan dilakukan.

**Tabel 3.1** Operasionalisasi Variabel

Variabel	Definisi	Satuan	Skala
(1)	(2)	(3)	(4)
Tahun ( $X_1$ )	Periode waktu revolusi bumi dalam kalender Masehi saat data penumpang tercatat, mencakup rentang tahun 2015 hingga 2025.	Tahun	-
Bulan ( $X_2$ )	Satuan waktu yang membagi satu tahun	Bulan	-

	menjadi 12 periode (Januari s.d. Desember) yang menunjukkan kapan aktivitas penerbangan terjadi.		
Total Kedatangan ( $X_3$ )	umlah penumpang angkutan udara yang mendarat (datang) di seluruh bandara komersial di wilayah provinsi Sumatera Selatan (seperti Bandara SMB II) dari bandara asal (baik domestik maupun internasional) dalam periode satu bulan.	Orang per Bulan	Rasio
Total Keberangkatan ( $X_4$ )	Jumlah penumpang angkutan udara yang terbang (berangkat) meninggalkan seluruh bandara komersial di wilayah provinsi Sumatera Selatan menuju bandara tujuan (baik domestik maupun internasional) dalam periode satu bulan.	Orang per Bulan	Rasio

Total Penumpang ( $X_5$ )	Akumulasi jumlah keseluruhan penumpang angkutan udara, yang merupakan penjumlahan dari total kedatangan dan total keberangkatan di wilayah Sumatera Selatan dalam periode satu bulan.	Orang Per Bulan	Rasio
Tanggal ( $X_6$ )	Variabel waktu yang dikonstruksi dari gabungan 'Tahun' dan 'Bulan' (biasanya diset ke tanggal 1 setiap bulannya, misal: 2015-01-01) sebagai indeks waktu ( <i>time index</i> ) untuk keperluan analisis <i>time series</i> .	Tanggal	-

Berdasarkan Tabel 3.1 Operasionalisasi Variabel, penelitian ini menggunakan dua jenis variabel, yaitu variabel utama dan variabel pendukung. Variabel utama dalam penelitian ini adalah Volume Penumpang Angkutan Udara (yang dirinci menjadi Total Kedatangan, Total Keberangkatan, dan Total Penumpang), yang merepresentasikan tingkat mobilitas masyarakat yang menggunakan moda transportasi udara di Sumatera Selatan. Variabel ini bersifat kuantitatif dengan satuan orang per bulan dan menggunakan skala rasio sehingga memungkinkan dilakukan analisis statistik serta simulasi pembangkitan data. Selain itu, penelitian ini juga menggunakan Elemen Waktu (Tahun, Bulan, dan Tanggal) sebagai variabel pendukung, yang berfungsi untuk mengidentifikasi periode kronologis saat data diobservasi. Variabel ini digunakan sebagai indeks waktu (*time index*) dengan cakupan data dari tahun 2015



hingga 2025. Penggunaan variabel ini memungkinkan peneliti untuk meninjau pola tren, fluktuasi musiman, serta distribusi data *time series* antar periode waktu.

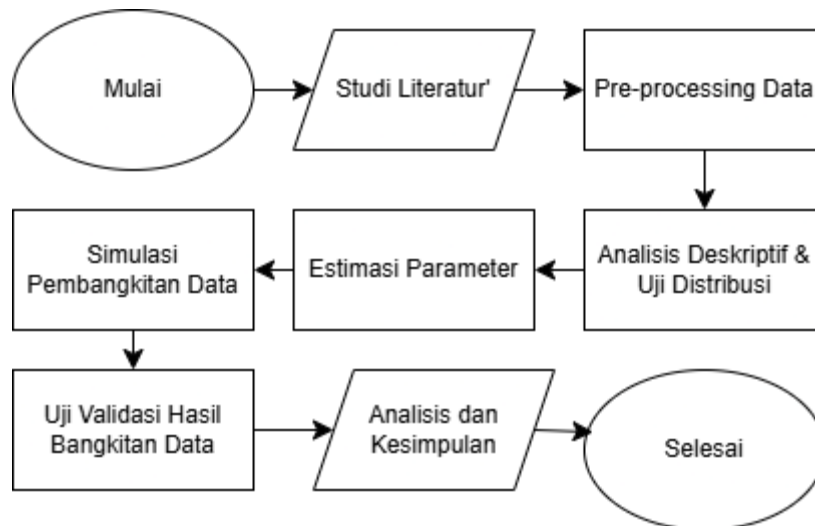
**Tabel 3.2** Data Penumpang Angkutan Udara Sumatera Selatan

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
2015	1	72171	62791	134962	2015-01-01
2015	2	169564	117399	286963	2015-02-01
2015	3	126843	125139	251982	2015-03-01
2015	4	129938	133238	263176	2015-04-01
2015	5	142324	139548	281872	2015-05-01
2015	6	134634	132166	266800	2015-06-01
...	...	...	...	...	...
2025	4	125175	142059	267234	2025-04-01
2025	5	103701	109938	213639	2025-04-01
2025	6	118576	115871	234447	2025-04-01
2025	7	127781	121020	248801	2025-04-01
2025	8	121788	118504	240292	2025-04-01
2025	9	119536	120654	240190	2025-04-01

Tabel 3.2 menyajikan struktur dataset penumpang angkutan udara yang digunakan sebagai basis dalam penelitian ini. Data tersebut terdiri dari enam variabel ( $X_1$  hingga  $X_6$ ) yang mencakup dimensi waktu dan volume penumpang. Variabel  $X_1$  (Tahun) dan  $X_2$  (Bulan) berfungsi sebagai penanda temporal yang menunjukkan periode terjadinya pergerakan penumpang. Variabel  $X_3$  merepresentasikan Total Penumpang Berangkat, yaitu jumlah penumpang yang meninggalkan bandara di Sumatera Selatan, sedangkan variabel  $X_4$  merepresentasikan Total Penumpang Datang, yaitu jumlah penumpang yang tiba di bandara tersebut. Penjumlahan dari kedua variabel ini membentuk variabel  $X_5$ , yaitu Total Penumpang, yang menjadi variabel respon utama dalam simulasi pembangkitan bilangan acak ini. Selain itu, terdapat variabel  $X_6$  (Tanggal) yang merupakan transformasi gabungan dari tahun dan bulan ke dalam format tanggal standar (*date object*) untuk memfasilitasi pengolahan data runtun waktu (*time series*) pada perangkat lunak R. Data yang ditampilkan mencakup rentang waktu pengamatan dari tahun 2015 hingga 2025.

### 3.4. Tahapan Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis untuk menjamin validitas hasil simulasi. Prosedur penelitian dimulai dari pengumpulan data hingga evaluasi kinerja algoritma. Rincian langkah-langkah penelitian adalah sebagai berikut:



**Gambar 3.1** Flowchart Penelitian

1. Tahap awal melibatkan studi literatur mengenai metode pembangkitan bilangan acak (*Inverse-Transform*, *Acceptance-Rejection*, dan *Box-Muller*) serta pengumpulan data sekunder. Data yang digunakan adalah data jumlah penumpang angkutan udara di Sumatera Selatan periode 2015–2025 yang diperoleh dari publikasi Badan Pusat Statistik (BPS).
2. Data mentah yang diperoleh disusun ulang sesuai struktur pada Tabel 3.2. Dilakukan pengecekan terhadap nilai yang hilang (*missing values*) dan penyesuaian format variabel waktu ( $X_1$  dan  $X_2$ ) menjadi format tanggal ( $X_6$ ) agar dapat dibaca sebagai objek *time series* dalam perangkat lunak R.
3. Dilakukan eksplorasi data menggunakan visualisasi (histogram dan plot *time series*) untuk memahami karakteristik data. Selanjutnya, dilakukan uji normalitas menggunakan metode *Shapiro-Wilk* untuk memverifikasi asumsi bahwa data penumpang mengikuti distribusi Normal.
4. Berdasarkan hasil uji distribusi, dilakukan estimasi parameter populasi berupa rata-rata ( $\mu$ ) dan standar deviasi ( $\sigma$ ) menggunakan metode *Maximum Likelihood Estimation* (MLE). Parameter ini digunakan sebagai input dalam fungsi pembangkit bilangan acak.
5. Tahap ini merupakan inti penelitian, di mana dilakukan pengembangan algoritma simulasi menggunakan bahasa pemrograman R. Tiga metode yang diimplementasikan adalah:
  - a. *Inverse-Transform Method*
  - b. *Acceptance-Rejection Method* (dengan fungsi proposal Cauchy)
  - c. *Transformation Method* (Box-Muller)
 Simulasi dilakukan dengan memvariasikan ukuran sampel ( $N = 100, 1.000, 10.000$ ) untuk menguji konsistensi metode.
6. Data hasil simulasi divalidasi menggunakan uji *Kolmogorov-Smirnov* (KS-Test) untuk melihat kesesuaian distribusinya dengan data riil. Selain itu, dilakukan perhitungan bias statistik (selisih *mean* dan *varians*) untuk mengukur tingkat kesalahan estimasi.
7. Hasil validasi dari ketiga metode dibandingkan untuk menentukan metode mana yang paling akurat dan efisien. Berdasarkan analisis tersebut, ditarik kesimpulan mengenai metode terbaik untuk mensimulasikan data penumpang angkutan udara.

## BAB IV

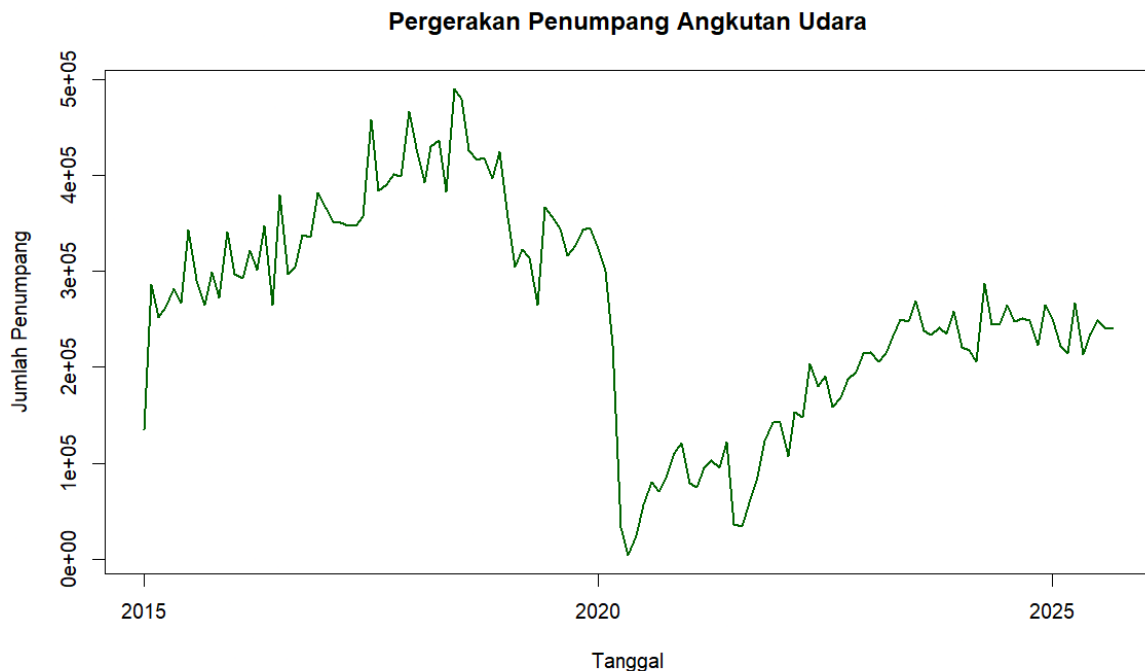
### HASIL DAN PEMBAHASAN

#### 4.1. Eksplorasi Data Penumpang Angkutan Udara

Sebelum melakukan pemodelan simulasi menggunakan metode *Inverse-Transform*, *Acceptance-Rejection*, dan *Box-Muller*, langkah pertama adalah memahami karakteristik data aktual. Data yang digunakan merupakan data runtun waktu (*time series*) jumlah total penumpang angkutan udara di Sumatera Selatan periode 2015 hingga 2025 yang bersumber dari Badan Pusat Statistik (BPS). Eksplorasi ini bertujuan untuk mengidentifikasi pola tren, musiman, serta bentuk distribusi data guna menentukan pendekatan simulasi yang tepat.

##### 4.1.1. Pola Pergerakan dan Dekomposisi Deret Waktu

Pemeriksaan deskriptif awal dilakukan dengan melihat komposisi deret waktu.

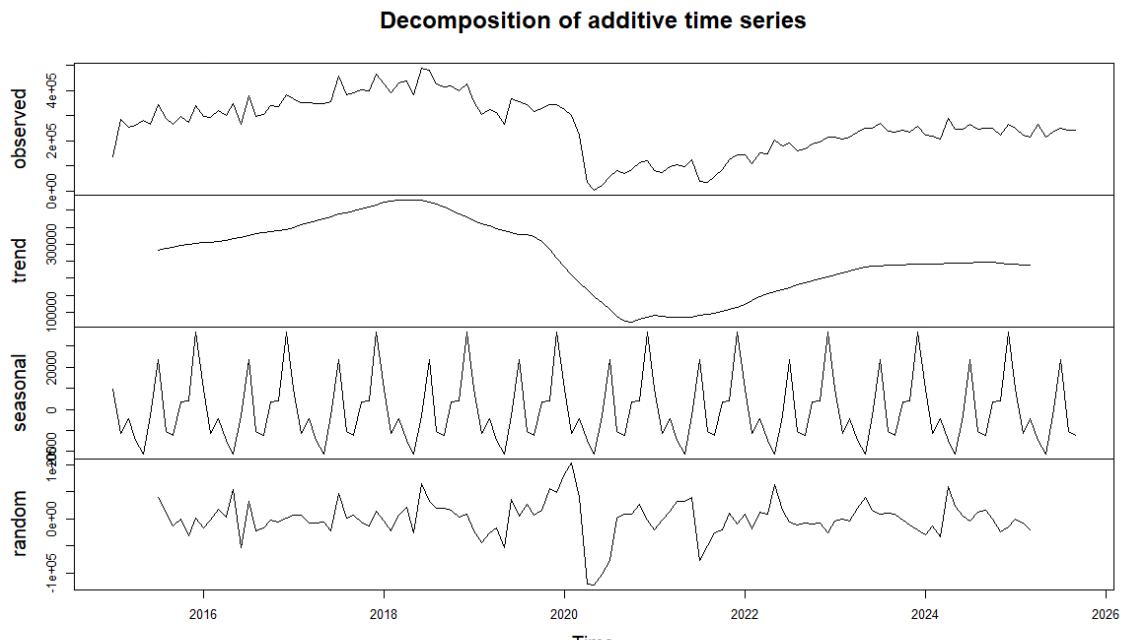


**Gambar 4.2** Pergerakan Penumpang Angkutan Udara Dalam Periode Tahun 2015 -2025

Berdasarkan visualisasi pergerakan penumpang (line chart) pada **Gambar 4.1**, terlihat dinamika yang sangat fluktuatif selama periode pengamatan. Secara umum, data dapat dibagi menjadi tiga fase utama:

1. Terlihat tren kenaikan jumlah penumpang yang stabil dengan puncak tertinggi (*peak*) terjadi sekitar tahun 2018-2019, di mana jumlah penumpang mencapai angka di atas 350.000 orang per bulan.
2. Terjadi penurunan struktural yang sangat tajam (*structural break*) pada awal tahun 2020. Grafik menunjukkan penurunan drastis hingga titik terendah mendekati 0, yang merefleksikan pembatasan mobilitas akibat pandemi COVID-19.

3. Mulai tahun 2021, terlihat tren positif pemulihan (*recovery*). Meskipun belum sepenuhnya kembali ke level puncak tahun 2018, data menunjukkan kenaikan bertahap dengan rata-rata di kisaran 250.000 hingga 300.000 penumpang.



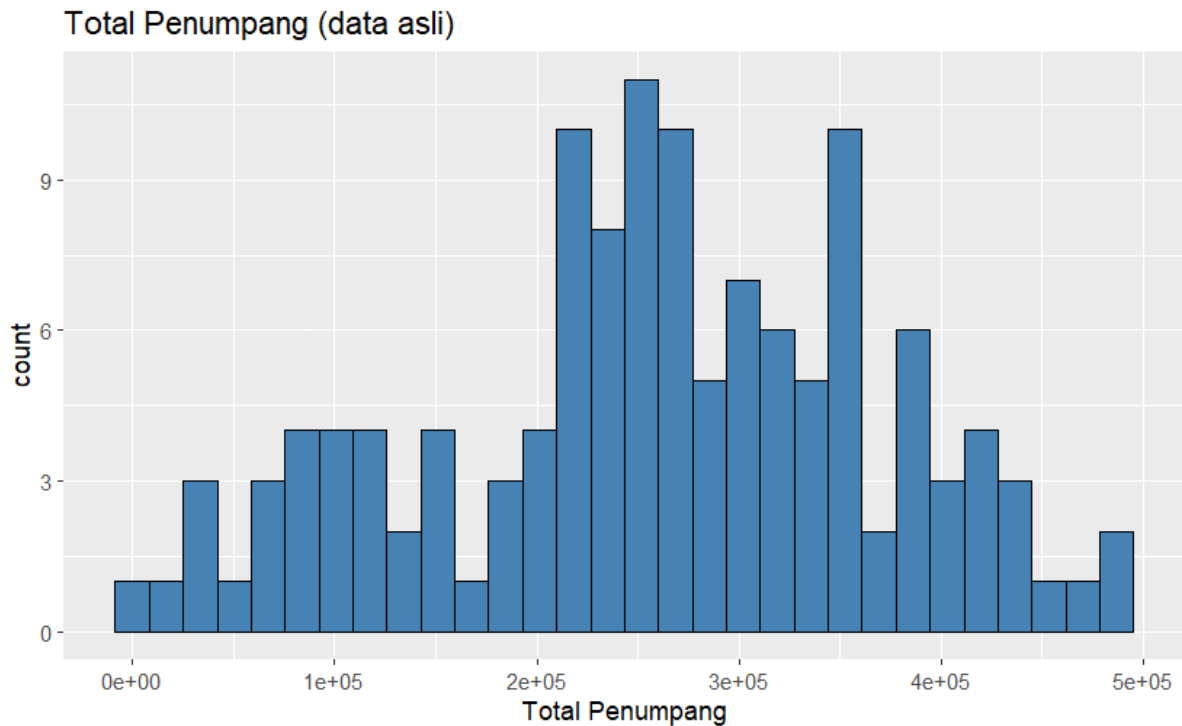
**Gambar 4.3** Dekomposisi Deret Waktu dari Data Total Penumpang Angkutan Udara

Analisis lebih lanjut menggunakan Dekomposisi Time Series (*Additive Model*) pada gambar 4.2 memperjelas komponen pembentuk data tersebut:

1. Komponen tren mengkonfirmasi penurunan tajam di tahun 2020 dan pemulihan lambat setelahnya.
2. Terdapat pola musiman yang kuat dan berulang setiap tahunnya. Pola gelombang yang konsisten pada komponen *seasonal* mengindikasikan adanya lonjakan penumpang pada bulan-bulan tertentu, kemungkinan besar terkait dengan periode libur hari raya (Lebaran) atau libur akhir tahun.
3. Komponen sisaan (*random/noise*) menunjukkan variabilitas yang tidak dapat dijelaskan oleh tren maupun musiman, yang cukup fluktuatif terutama pada periode transisi pandemi.

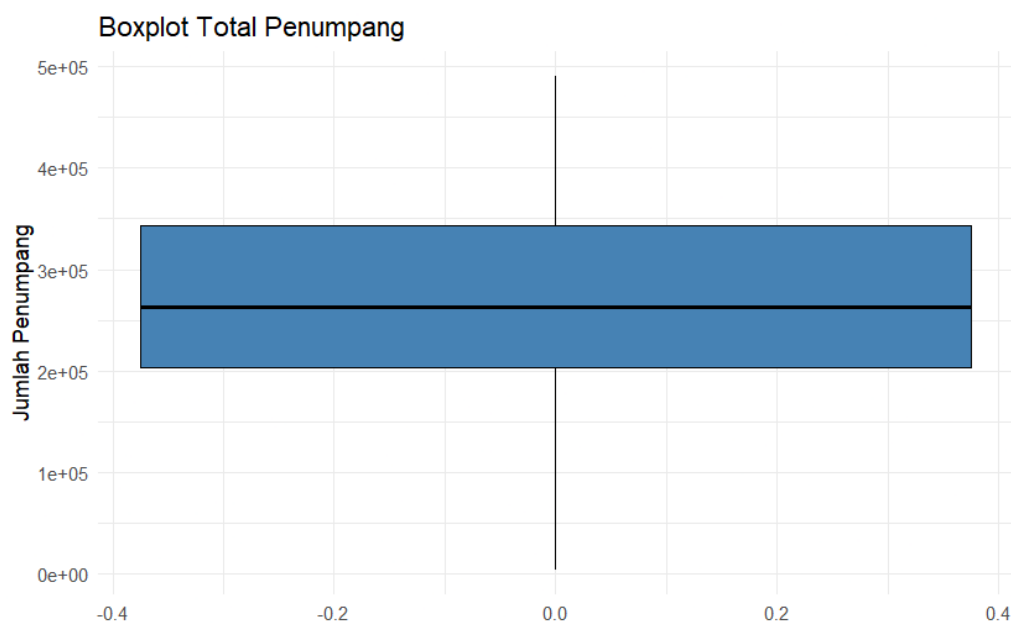
#### 4.1.2. Karakteristik Distribusi Data

Untuk keperluan pembangkitan bilangan acak, pemahaman mengenai distribusi data sangat krusial. Visualisasi histogram dan boxplot digunakan untuk melihat sebaran data secara keseluruhan.



**Gambar 4.3** Histogram Data Historis Total Penumpang Angkutan Udara di Sumatera Selatan

Berdasarkan histogram pada gambar 4.3 data total penumpang, terlihat bahwa distribusi data cenderung memusat di tengah namun tidak sepenuhnya simetris sempurna. Frekuensi tertinggi (modus) berada pada interval 200.000 – 300.000 penumpang. Namun, terdapat pemanjangan ekor ke arah kiri (*left-skewed* atau *negative skewness*). Hal ini disebabkan oleh adanya nilai-nilai observasi yang sangat rendah (mendekati 0) yang terjadi selama periode pandemi COVID-19. Meskipun demikian, bentuk umum histogram masih menyerupai kurva lonceng (*bell-shaped*), yang memberikan indikasi awal bahwa pendekatan distribusi Normal mungkin dapat



**Gambar 4.4** Boxplot Total Penumpang Angkutan Udara Sumatera Selatan Tahun 2015-2025

digunakan dengan catatan adanya varians yang besar. Hal ini diperkuat oleh visualisasi boxplot. Boxplot menunjukkan:

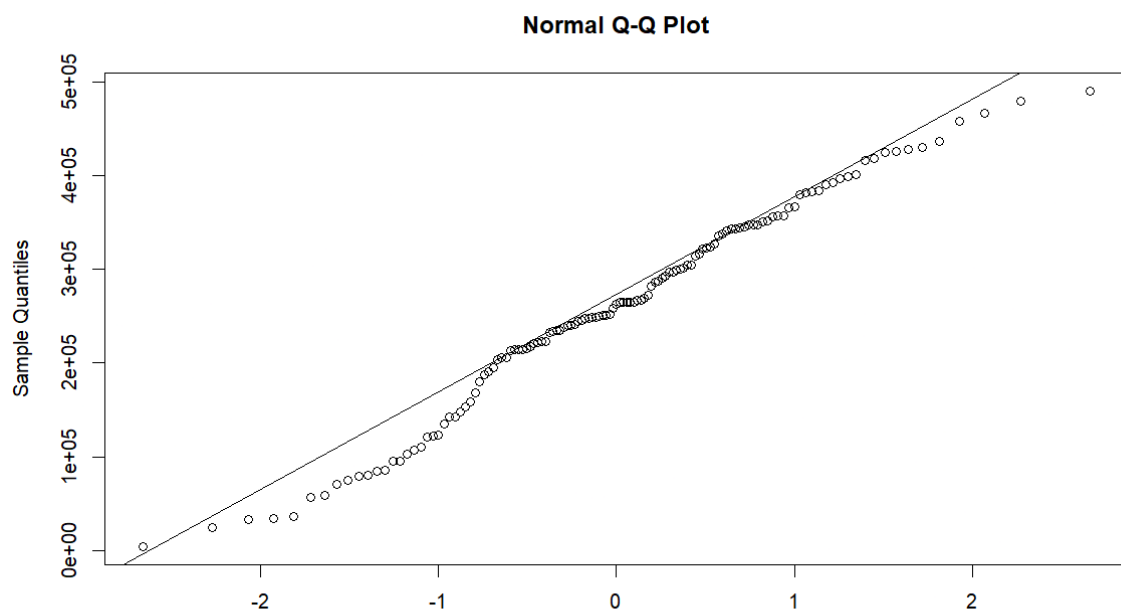
1. Garis tengah boxplot berada di kisaran angka 270.000-an, yang membagi data menjadi dua bagian sama besar.
2. Kotak (Interquartile Range/IQR) cukup lebar, menandakan variasi jumlah penumpang antar bulan cukup tinggi.
3. Boxplot tidak mendeteksi adanya *outlier* (pencilan) ekstrem di luar *whisker* atas maupun bawah. *Whisker* bawah yang memanjang hingga mendekati 0 menunjukkan bahwa nilai rendah saat pandemi, meskipun ekstrem secara fenomena, secara statistik masih dianggap berada dalam rentang sebaran data

#### 4.2. Uji Asumsi Distribusi dan Estimasi Parameter

Setelah mengetahui karakteristik deskriptif data, langkah krusial berikutnya adalah memvalidasi asumsi distribusi. Ketiga metode simulasi yang akan digunakan (*Inverse-Transform*, *Acceptance-Rejection*, dan *Box-Muller*) pada penelitian ini dirancang untuk membangkitkan bilangan acak yang mengikuti distribusi Normal Standar. Oleh karena itu, perlu dibuktikan apakah data aktual penumpang angkutan udara ini memenuhi asumsi kenormalan tersebut.

##### 4.2.1. Uji Normalitas Data

Pemeriksaan normalitas dilakukan melalui dua pendekatan: pemeriksaan visual menggunakan *Normal Q-Q Plot* dan pengujian formal menggunakan *Shapiro-Wilk Normality Test*. Gambar 4.5 memperlihatkan sebaran kuantil data sampel terhadap kuantil teoritis distribusi normal.



Gambar 4.5 Normal Q-Q Plot dari Data Total Penumpang Angkutan Udara

Secara visual, titik-titik data (lingkaran) cenderung mengikuti garis diagonal lurus, terutama pada bagian tengah distribusi. Hal ini mengindikasikan bahwa sebagian besar data selaras dengan distribusi normal. Meskipun demikian, terlihat adanya sedikit deviasi pada bagian ekor bawah (*lower tail*), di mana titik-titik data berada sedikit di bawah garis referensi.

Penyimpangan di ekor kiri ini konsisten dengan temuan pada histogram sebelumnya yang menunjukkan adanya *negative skewness* akibat penurunan jumlah penumpang yang ekstrim saat pandemi. Untuk memastikan interpretasi visual tersebut, dilakukan uji hipotesis formal menggunakan uji Shapiro-Wilk dengan hipotesis sebagai berikut:

$H_0$ : Data berdistribusi Normal.

$H_1$ : Data tidak berdistribusi Normal.

Berdasarkan hasil komputasi menggunakan R, diperoleh nilai statistik uji  $W = 0.9814$  dengan  $p - value = 0.07387$ . Dengan menggunakan tingkat signifikansi ( $\alpha$ ) sebesar 0.05, didapatkan bahwa nilai  $p - value > 0.05$  ( $0.07387 > 0.05$ ). Keputusan pengujian adalah gagal tolak  $H_0$ . Artinya, tidak cukup bukti statistik untuk menyatakan bahwa data menyimpang dari distribusi normal. Dengan demikian, dapat disimpulkan bahwa data jumlah penumpang angkutan udara di Sumatera Selatan berdistribusi normal. Kesimpulan ini menjadi landasan yang valid untuk melanjutkan proses simulasi menggunakan parameter distribusi normal.

#### 4.2.2. Estimasi Parameter (Maximum Likelihood Estimation)

Setelah asumsi normalitas terpenuhi, langkah selanjutnya adalah mengestimasi parameter populasi. Mengingat distribusi normal ditentukan oleh dua parameter utama, yaitu rata-rata ( $\mu$ ) dan simpangan baku ( $\sigma$ ), maka dilakukan estimasi parameter menggunakan metode *Maximum Likelihood Estimation* (MLE) menggunakan fungsi *fitdist* di R.

Output:

```
mean
258230.8
sd
110611.3
```

Nilai estimasi ini menunjukkan bahwa rata-rata jumlah penumpang per bulan selama periode 2015-2025 adalah sekitar 258 ribu orang. Namun, nilai standar deviasi yang cukup besar (sekitar 42% dari rata-rata) mengonfirmasi volatilitas data yang tinggi, yang sejalan dengan temuan pada analisis *time series* sebelumnya di mana terjadi guncangan hebat akibat pandemi.

#### 4.3. Implementasi Simulasi Pembangkit Bilangan Acak

Pada bagian ini, dilakukan pembangkitan data tiruan sebanyak 100, 1000, dan 10000 observasi menggunakan tiga pendekatan algoritma yang berbeda. Tujuannya adalah membandingkan seberapa baik masing-masing metode dan jumlah observasi yang dibangkitkan dalam mereplikasi karakteristik data penumpang BPS Sumatera Selatan.

##### 4.3.1 Metode *Inverse-Transform*

Metode simulasi pertama yang diterapkan adalah Inverse-Transform Method. Metode ini didasarkan pada prinsip Probability Integral Transform (PIT), di mana jika  $U$  adalah variabel acak berdistribusi Uniform (0,1), maka  $X = F^{-1}(U)$  akan memiliki fungsi distribusi kumulatif  $F$ .

### a. Algoritma dan Implementasi

Untuk membangkitkan data penumpang yang diasumsikan berdistribusi Normal  $(\mu, \sigma)$ , algoritma yang diimplementasikan di R ada pada algoritma 4.1:

---

#### Algoritma 4.1 Metode *Inverse-Transform*

---

**Input:** Ukuran sampel  $n$ , Parameter Rata-rata  $\mu$ , Parameter Standar Deviasi  $\sigma$

**Output:** Vektor data simulasi  $X$  yang berdistribusi  $N(\mu, \sigma)$

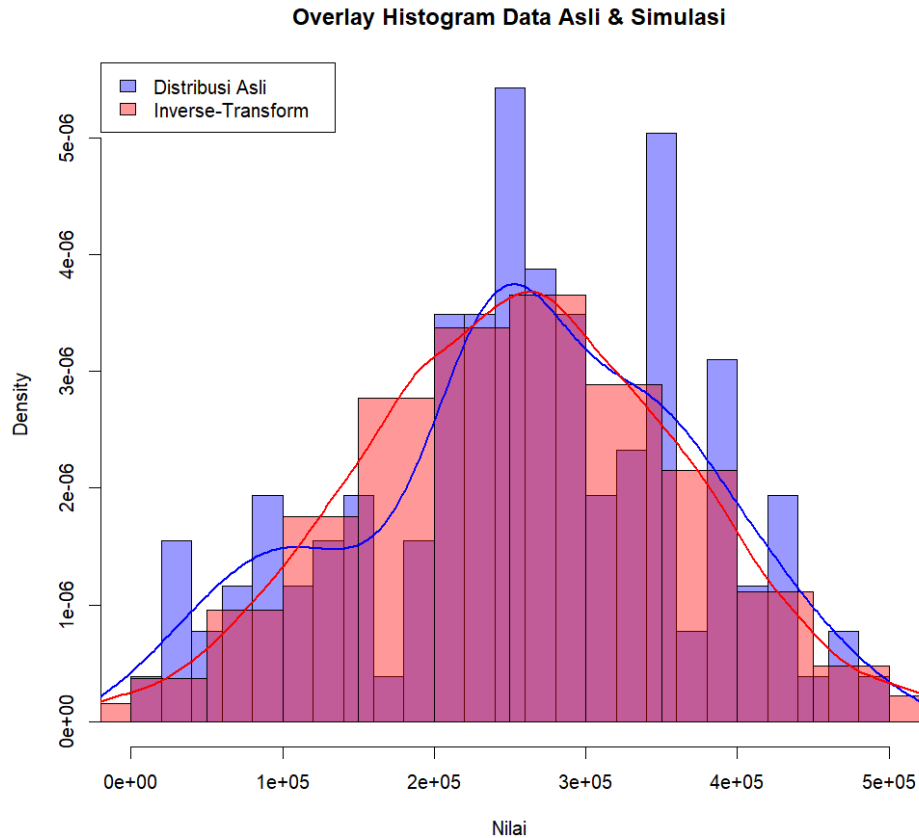
---

1. **Function** InverseTransformSimulation( $n, \hat{\mu}, \hat{\sigma}$ )
  2. Inisialisasi vektor  $U$  dan  $X$  dengan panjang  $n$
  3. Bangkitkan  $U \sim \text{Uniform}(0,1)$  sebanyak  $n$
  4. **For**  $i \leftarrow 1$  **to**  $n$  **do**
  5.     Hitung  $X[i] \leftarrow \Phi^{-1}(U[i]) \times \hat{\sigma} + \hat{\mu}$
  6. **End For**
  7. **Return**  $X$
  8. **End Function**
- 

Implementasi komputasi metode ini dalam lingkungan pemrograman R diawali dengan inisialisasi *seed*, yang bertujuan untuk menjamin konsistensi dan reproduksibilitas hasil simulasi pada setiap eksekusi kode. Proses inti algoritma dilakukan dengan membangkitkan vektor bilangan acak  $U$  yang mengikuti distribusi seragam standar atau  $U \sim \text{Uniform}(0,1)$ . Nilai-nilai probabilitas  $U$  tersebut kemudian dipetakan ke dalam domain data riil melalui fungsi invers distribusi kumulatif (*quantile function*) menggunakan perintah `qnorm`. Dalam proses transformasi ini, parameter  $\hat{\mu}$  dan  $\hat{\sigma}$  yang telah diestimasi sebelumnya dari data BPS diaplikasikan secara langsung, sehingga bilangan acak yang dihasilkan memiliki karakteristik sebaran yang sesuai dengan populasi target.

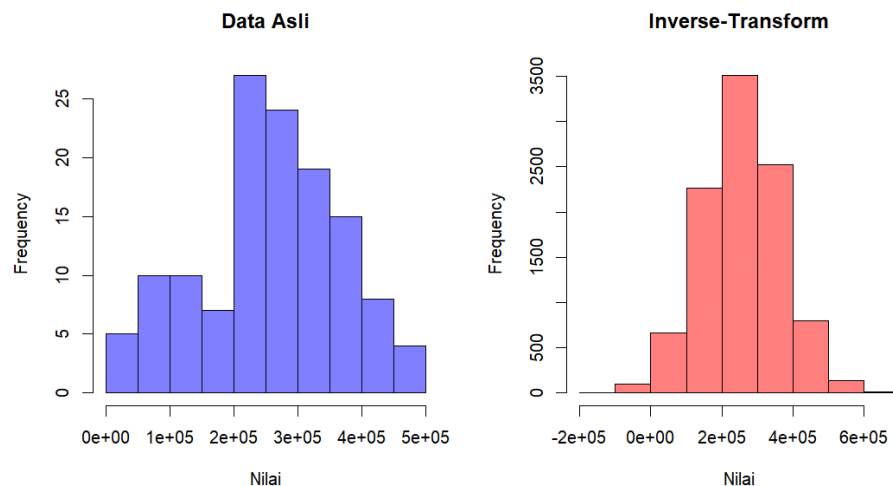
### b. Analisis Visual





**Gambar 4.6** Overlay Histogram Data Asli & Simulasi

Berdasarkan Gambar 4.6, terlihat kurva densitas hasil simulasi mampu mengikuti pola distribusi data asli dengan cukup baik. Simulasi berhasil menangkap karakteristik utama data, yaitu pemusatan data (*central tendency*) di angka 200.000 – 300.000 penumpang.



**Gambar 4.7** Perbandingan Histogram Data Asli & Inverse-Transform

Pada Gambar 4.7, terlihat bahwa histogram hasil simulasi memiliki bentuk lonceng yang lebih sempurna dibandingkan data asli. Hal ini wajar karena data simulasi dibangkitkan dari distribusi teoritis yang ideal, sedangkan data asli memiliki *noise* dan ketidakaturan akibat

faktor eksternal (seperti pandemi). Namun, secara umum, *range* dan penyebaran data simulasi sangat mirip dengan data aktual.

### c. Validasi Statistik (Uji Kolmogorov-Smirnov)

Untuk menguji kesesuaian distribusi secara formal, dilakukan *Two-Sample Kolmogorov-Smirnov Test* antara data asli dengan data hasil simulasi.

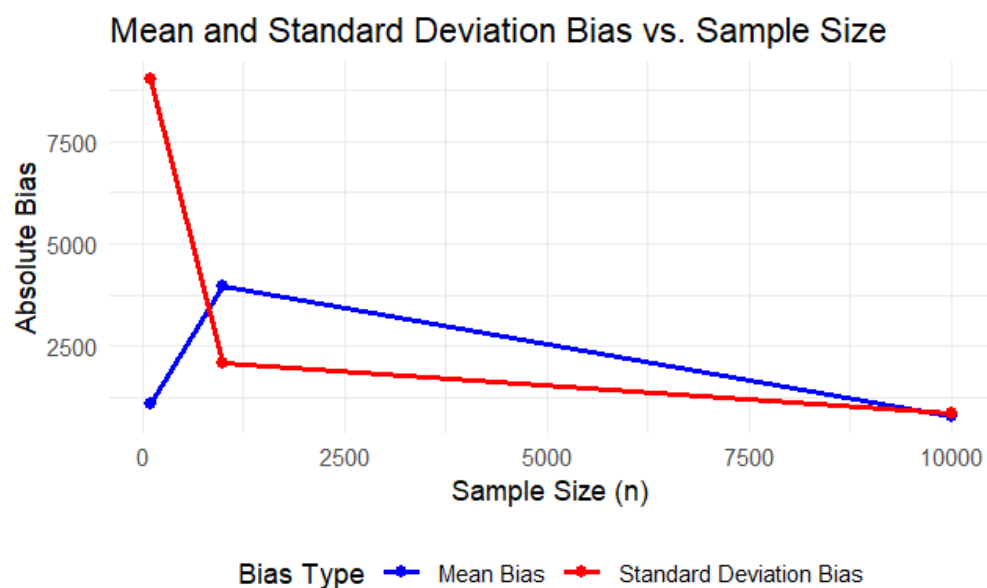
```
Asymptotic two-sample Kolmogorov-Smirnov test

data:  df$Total_Penumpang and x_inv
D = 0.073482, p-value = 0.4973
alternative hypothesis: two-sided
```

Karena nilai  $p - value > 0.05$ , maka keputusan pengujian adalah gagal tolak  $H_0$ . Hal ini membuktikan secara statistik bahwa tidak ada perbedaan signifikan antara distribusi data penumpang asli dengan data yang dibangkitkan menggunakan metode *Inverse-Transform*. Metode ini valid untuk digunakan.

### d. Analisis Konsistensi Ukuran Sampel

Selain validasi distribusi, dilakukan pula pengujian terhadap konsistensi metode dengan memvariasikan ukuran sampel ( $N = 100, 1000, 10000$ ).



**Gambar 4.8** Perbandingan Bias Mean dan Standar Deviasi Berdasarkan Nilai N

Berdasarkan gambar 4.8, terlihat pola konvergensi yang jelas, pada  $n = 100$ , bias (kesalahan estimasi) standar deviasi sangat tinggi, mencapai angka 9033. Pada  $n = 1000$ , bias mulai menurun secara signifikan. Pada  $n = 10000$ , bias mencapai titik terendah, di mana selisih mean hanya sebesar 771 dan selisih standar deviasi sebesar 884. Fenomena ini mengkonfirmasi Hukum Bilangan Besar (Law of Large Numbers), di mana semakin besar ukuran sampel simulasi, maka estimator parameter akan semakin mendekati parameter populasi aslinya (bias

mendekati nol). Oleh karena itu, untuk representasi data terbaik, penggunaan  $n$  yang besar ( $n = 10.000$ ) pada metode *Inverse-Transform* sangat disarankan karena memberikan akurasi parameter yang paling tinggi.

### 4.3.2 Metode *Acceptance-Rejection*

Metode kedua yang diterapkan adalah *Acceptance-Rejection Method*. Berbeda dengan metode invers yang memerlukan fungsi invers CDF, metode ini bekerja dengan cara membangkitkan bilangan acak dari distribusi proposal  $g(x)$  yang lebih sederhana, kemudian menerima atau menolak bilangan tersebut berdasarkan rasio densitasnya terhadap distribusi target  $f(x)$ .

#### a. Algoritma dan Interpretasi

---

Algoritma 4.2: Metode *Acceptance-Rejection*

---

**Input:** Jumlah sampel  $n$ , Rata-rata  $\mu$ , Standar Deviasi  $\sigma$

**Output:** Vektor data simulasi  $X$

---

```

1. Function AcceptanceRejection( $n, \mu, \sigma$ )
2.    $count \leftarrow 0$ 
3.   Inisialisasi array  $X$  kosong dengan ukuran  $n$ 
4.   While  $count < n$  do
5.     Bangkitkan kandidat  $Y$  dari distribusi Cauchy standar
6.     Bangkitkan  $U \sim Uniform(0,1)$ 
7.     Hitung densitas target  $f(Y)$  (Normal)
8.     Hitung densitas proposal  $g(Y)$  (Cauchy)
9.     If  $U \leq \frac{f(Y)}{c \cdot g(Y)}$  then
10.       $count \leftarrow count + 1$ 
11.       $X[count] \leftarrow Y$ 
12.    End If
13.  End While
14.   $X \leftarrow \mu + (X \times \sigma)$ 
15.  Return  $X$ 
16. End Function

```

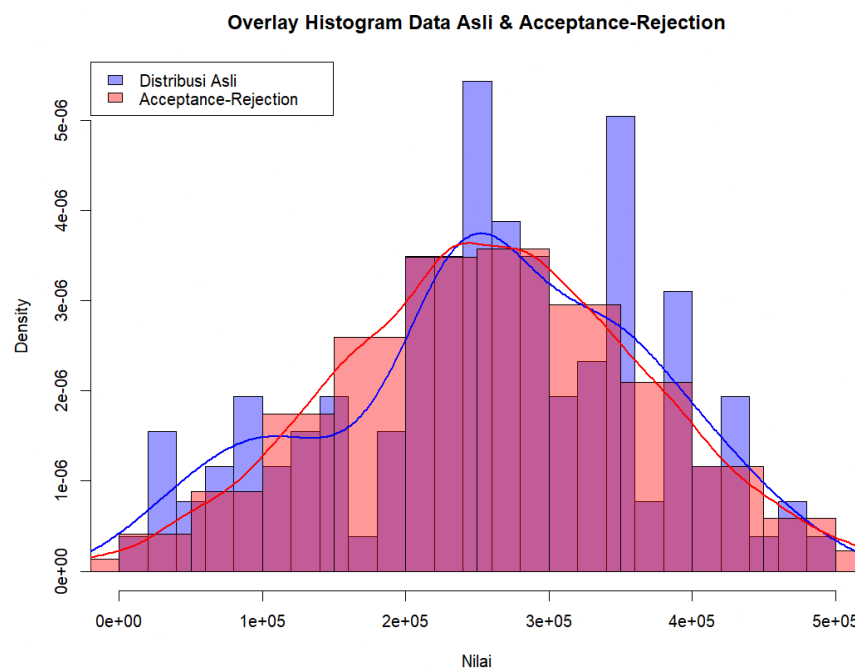
---

Dalam implementasi pada algoritma 4.2, distribusi target yang dituju adalah normal standar, yang kemudian akan ditransformasi ke skala data penumpang. Algoritma menggunakan distribusi Cauchy sebagai fungsi proposal ( $g(x)$ ) karena memiliki karakteristik heavy-tailed yang mampu mendominasi distribusi Normal pada bagian ekor. Konstanta pengali ditetapkan sebesar  $c = 1.5$  untuk memastikan fungsi selubung  $c \cdot g(x)$  selalu berada di atas  $f(x)$ . Proses

iterasi dilakukan dengan membangkitkan kandidat  $Y \sim \text{Cauchy}(0,1)$  dan bilangan acak seragam  $U$ . Jika kondisi  $U \leq \frac{f(Y)}{c \cdot g(Y)}$  terpenuhi, maka  $Y$  diterima sebagai bilangan acak normal standar. Terakhir, nilai yang diterima dikonversi ke satuan data asli menggunakan persamaan linear  $X = \hat{\mu} + \hat{\sigma}Z$ .

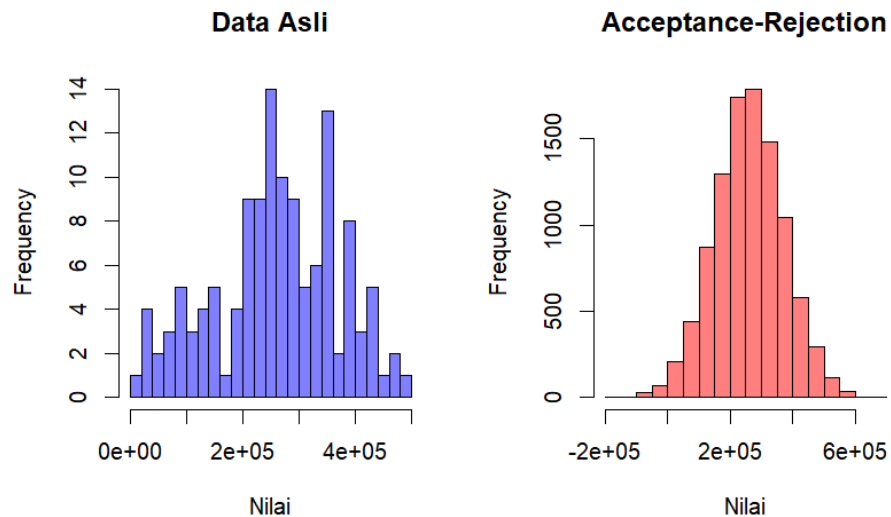
## b. Analisis Visual

Visualisasi pada Gambar 4.9 menunjukkan hasil simulasi yang sangat presisi. Kurva densitas dari data yang dibangkitkan (garis merah) menempel ketat pada pola distribusi data asli. Metode *Acceptance-Rejection* dengan proposal Cauchy ini terbukti mampu mereplikasi bentuk lonceng (*bell-shaped*) data penumpang dengan akurat.



**Gambar 4.9** Overlay Histogram Data Asli & Acceptance-Rejection

Pada Gambar 4.10, terlihat bahwa histogram hasil simulasi memiliki struktur yang lebih halus dan simetris dibandingkan data asli yang cenderung memiliki gap atau ketidakteraturan pada frekuensi tertentu. Namun, secara rentang nilai (range), data simulasi mencakup interval yang sama dengan data BPS (sekitar 0 hingga 500.000 penumpang), menunjukkan bahwa variabilitas data berhasil ditangkap dengan baik.



**Gambar 4.10** Perbandingan Histogram Data Asli dan Acceptance-Rejection

### c. Validasi Statistik (Uji Kolmogorov-Smirnov)

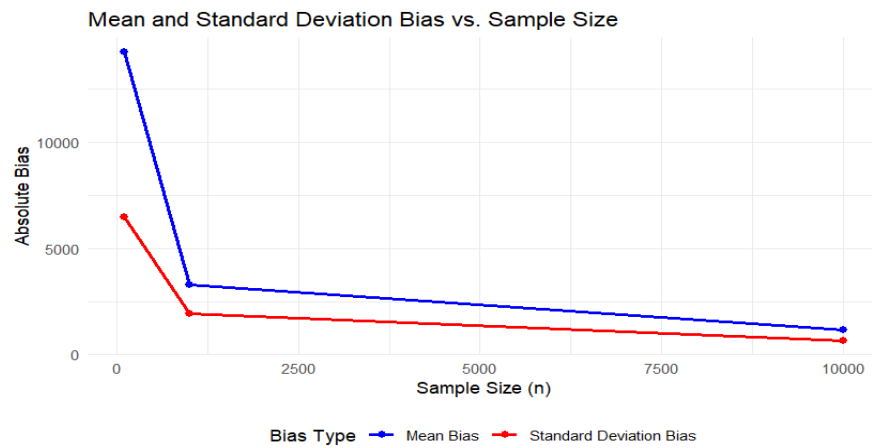
Untuk menguji kesesuaian distribusi secara formal, dilakukan *Two-Sample Kolmogorov-Smirnov Test* antara data asli dengan data hasil simulasi.

```
Asymptotic two-sample Kolmogorov-Smirnov test

data:  df$Total_Penumpang and x_ar
D = 0.061982, p-value = 0.7121
alternative hypothesis: two-sided
```

Nilai p-value (0.7121) ini bahkan lebih tinggi dibandingkan metode *Inverse-Transform* sebelumnya (0.49), yang mengindikasikan bahwa secara statistik, data yang dihasilkan oleh metode *Acceptance-Rejection* ini memiliki tingkat kesesuaian distribusi yang sangat kuat terhadap data riil. Dengan  $p > 0.05$ , keputusan pengujian adalah gagal tolak  $H_0$ , sehingga metode ini dinyatakan valid.

### d. Analisis Konsistensi Ukuran Sampel



**Gambar 4.11** Perbandingan Bias Mean dan Standar Deviasi Berdasarkan Nilai N

Selain validasi distribusi, dilakukan pula pengujian terhadap konsistensi metode dengan memvariasikan ukuran sampel ( $N = \{100, 1000, 10000\}$ ). Pada gambar 4.11 sampel kecil ( $n = 100$ ), bias parameter Mean sangat besar, mencapai 14.270 poin. Seiring peningkatan sampel ke  $n = 1000$  dan  $n = 10000$ , bias menurun drastis secara eksponensial. Pada  $n = 10.000$ , bias Mean tersisa di angka 1.150 dan bias Standar Deviasi di angka 645. Meskipun metode ini memiliki biaya komputasi yang lebih mahal karena adanya proses penolakan data (*rejection*), hasil validasi menunjukkan bahwa metode ini memberikan akurasi distribusi yang sangat tinggi ketika ukuran sampel memadai.

### 4.3.3 Transformation Method (Box-Muller)

Metode ketiga sekaligus terakhir yang diimplementasikan adalah Box-Muller Transformation. Berbeda dengan dua metode sebelumnya yang bekerja pada satu dimensi, metode ini memanfaatkan sifat geometri koordinat polar untuk membangkitkan sepasang bilangan acak Normal standar yang saling bebas (independent) dari dua bilangan acak Uniform.

#### a. Algoritma dan Implementasi

---

Algoritma 4.3: Metode *Box-Muller*

---

**Input:** Jumlah sampel  $n$ , Parameter  $\mu, \sigma$

**Output:** Vektor data simulasi  $X$

---

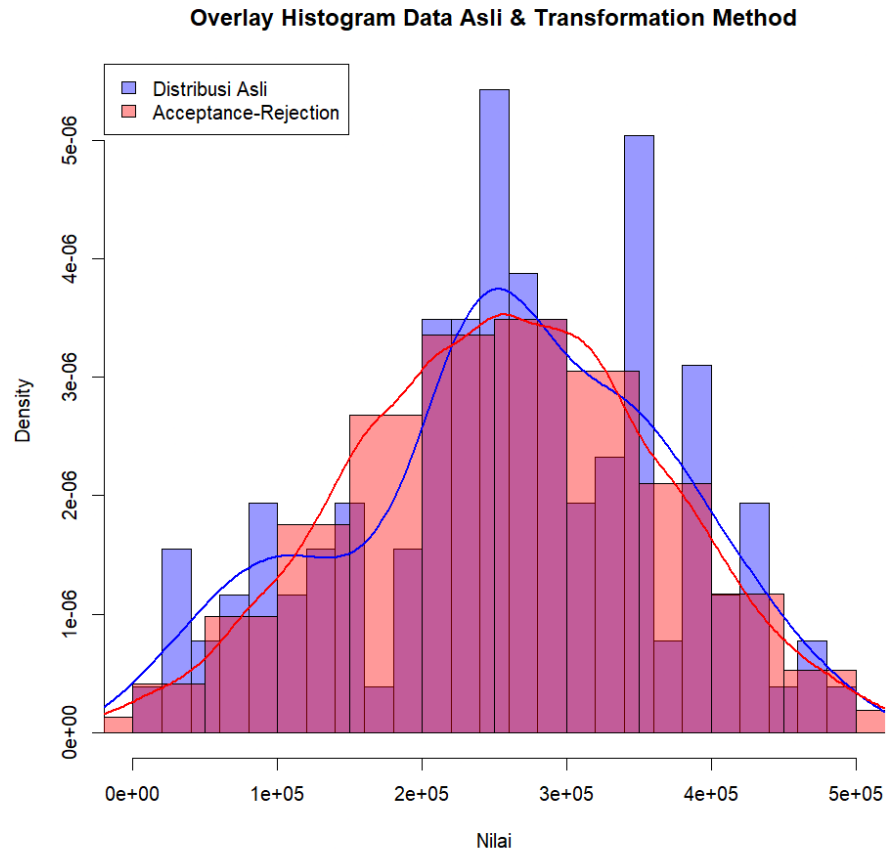
1. **Function** BoxMullerTransform( $n$ )
  2.  $m \leftarrow \lfloor n / 2 \rfloor$  (Setengah dari ukuran sampel)
  3. Bangkitkan  $U_1 \sim \text{Uniform}(0,1)$  sebanyak  $m$
  4. Bangkitkan  $U_2 \sim \text{Uniform}(0,1)$  sebanyak  $m$
  5. # Transformasi Polar ke Kartesius
  6.  $R \leftarrow \sqrt{-2 \ln(U_1)}$
  7.  $\theta \leftarrow 2\pi U_2$
  8.  $Z_1 \leftarrow R \cos(\theta)$
-

- 
9.  $Z_2 \leftarrow R \sin(\theta)$
  10. Gabungkan  $Z \leftarrow \{Z_1, Z_2\}$
  11. Potong  $Z$  agar panjangnya tepat  $n$
  12.  $X \leftarrow \mu + (Z \times \sigma)$
  13. **Return**  $X$
  14. **End Function**
- 

Algoritma Box-Muller pada algoritma 4.3 diimplementasikan dengan membangkitkan dua set bilangan acak seragam  $U_1$  dan  $U_2$  dari distribusi Uniform(0,1). Ide dasarnya adalah mentransformasi koordinat polar  $(R, \theta)$  menjadi koordinat Kartesius  $(Z_1, Z_2)$ . Dalam kode R yang disusun, variabel  $R$  dihitung menggunakan rumus  $R = \sqrt{-2 \ln U_1}$ , sedangkan sudut  $\theta$  dihitung dengan  $\theta = 2\pi U_2$ . Dari sini, diperoleh dua variat normal standar sekaligus melalui persamaan  $Z_1 = R \cos(\theta)$  dan  $Z_2 = R \sin(\theta)$ . Karena algoritma ini menghasilkan dua output dalam satu iterasi, efisiensi pembangkitan data menjadi lebih tinggi dibandingkan metode *Acceptance-Rejection* yang membuang data. Terakhir, nilai  $Z$  yang dihasilkan disesuaikan (diskalakan) dengan parameter  $\hat{\mu}$  dan  $\hat{\sigma}$  agar sesuai dengan karakteristik data penumpang BPS.

#### **b. Analisis Visual**

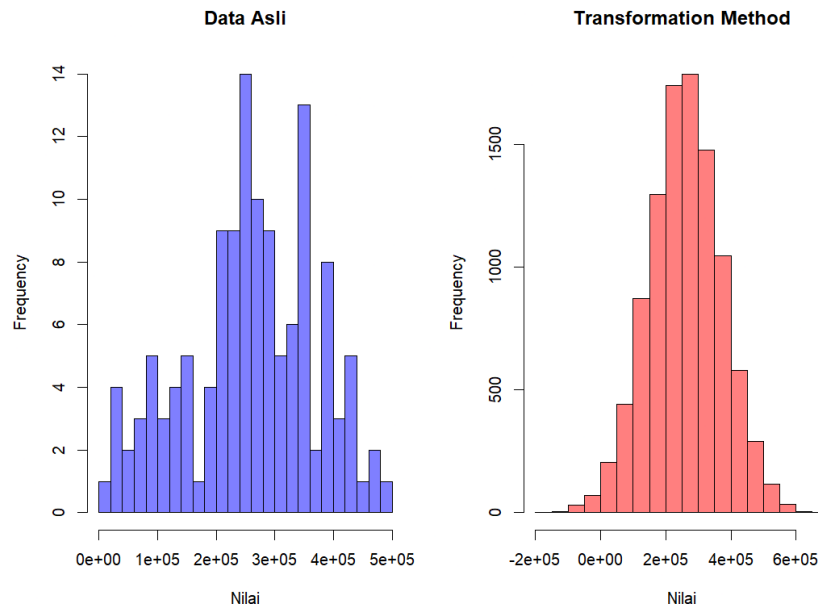
Evaluasi visual terhadap kinerja metode Box-Muller dilakukan melalui dua pendekatan grafis, yaitu histogram tumpang-tindih (*overlay*) dan plot perbandingan berdampingan (*side-by-side*).



**Gambar 4.12** Overlay Histogram Data Asli & Transformation Method

Berdasarkan Gambar 4.12, terlihat bahwa kurva densitas data hasil simulasi (garis merah) memiliki kesesuaian yang tinggi dengan histogram data asli (batang biru). Karena metode Box-Muller menggunakan transformasi eksak dari koordinat polar, sebaran data yang dihasilkan cenderung membentuk kurva lonceng (*bell-shaped*) yang sangat mulus dan simetris, menutupi area densitas utama data penumpang pada rentang 200.000 hingga 300.000 penumpang.





**Gambar 4.13** Perbandingan Histogram Data Asli dengan Transformation Method

Selanjutnya, pada Gambar 4.13, komparasi struktur data terlihat lebih jelas. Histogram Box-Muller menampilkan karakteristik distribusi Normal teoretis yang "bersih", namun tetap mempertahankan rentang nilai (*range*) dan variabilitas yang setara dengan data aktual BPS. Kemampuan metode ini dalam mereplikasi parameter lokasi ( $\mu$ ) dan skala ( $\sigma$ ) data asli terlihat dari posisi puncak histogram dan lebar sebarannya yang identik dengan data referensi.

### c. Validasi Statistik (Uji Kolmogorov-Smirnov)

Untuk menguji kesesuaian distribusi secara formal, dilakukan *Two-Sample Kolmogorov-Smirnov Test* antara data asli dengan data hasil simulasi.

```
Asymptotic two-sample Kolmogorov-Smirnov test
```

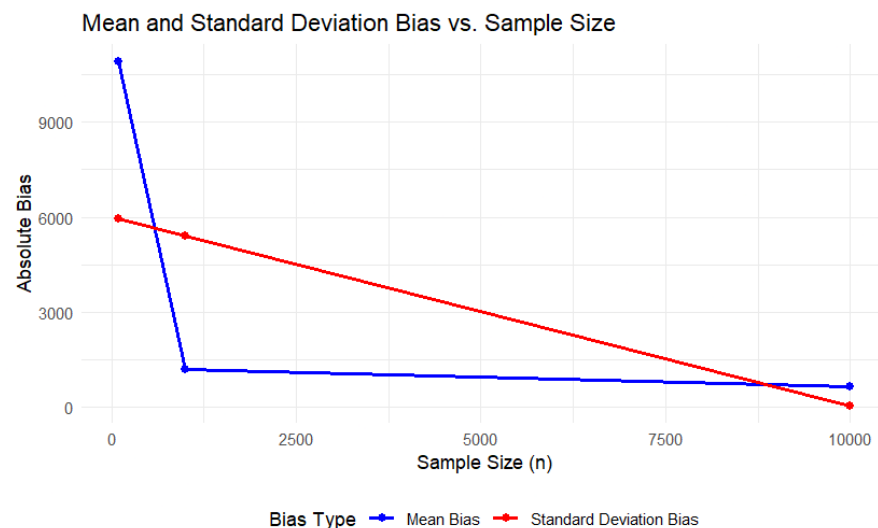
```
data: df$Total_Penumpang and x_box
```

```
D = 0.077482, p-value = 0.429
```

```
alternative hypothesis: two-sided
```

Hasil pengujian menunjukkan nilai  $p - value > 0.05$ . Dengan demikian, keputusan uji adalah gagal tolak  $H_0$ . Hal ini mengonfirmasi bahwa data yang dibangkitkan melalui transformasi Box-Muller memiliki distribusi yang konsisten dengan data penumpang angkutan udara Sumatera Selatan secara statistik.

### d. Analisis Konsistensi Ukuran Sampel



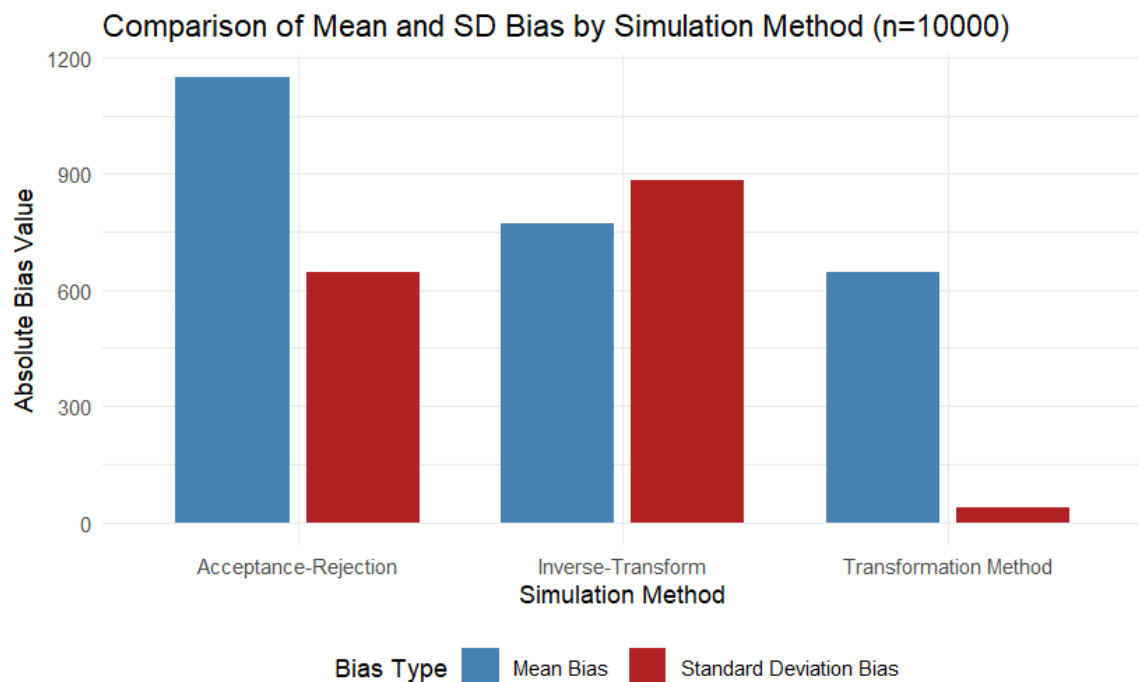
**Gambar 4.14** Perbandingan Nilai Bias Mean dan Standar Deviasi Berdasarkan N

Analisis sensitivitas ukuran sampel pada metode Box-Muller menunjukkan fenomena konvergensi yang bertahap, sebagaimana terlihat pada Gambar 4.14. Pada sampel kecil ( $n = 100$ ), tingkat bias masih tergolong sangat tinggi, di mana selisih rata-rata mencapai 10.909 dan bias standar deviasi sebesar 5.944. Hal ini wajar mengingat jumlah data yang minim belum mampu merepresentasikan populasi secara stabil. Pada sampel menengah ( $n = 1.000$ ), terjadi penurunan drastis pada bias rata-rata yang menyusut menjadi 1.209. Namun, menariknya, bias pada standar deviasi masih relatif tinggi di angka 5.395. Ini mengindikasikan bahwa pada ukuran sampel ini, metode sudah cukup baik mengestimasi pusat data, namun belum sepenuhnya presisi dalam menangkap varians/sebarannya. Pada sampel besar ( $n = 10.000$ ), akurasi mencapai titik optimalnya. *Mean Bias* turun kembali menjadi 647, dan yang paling signifikan adalah *Standard Deviation Bias* yang anjlok menjadi hanya 38,98. Penurunan bias yang drastis pada tahap akhir ini membuktikan konsistensi metode Box-Muller, di mana estimator parameter menjadi sangat presisi ketika ukuran sampel diperbesar hingga mendekati ukuran populasi yang memadai.

#### 4.4. Evaluasi dan Perbandingan Kinerja Metode

Bagian ini merupakan sintesis akhir yang bertujuan untuk mengevaluasi kinerja ketiga metode simulasi (*Inverse-Transform*, *Acceptance-Rejection*, dan *Transformation Method*) secara komprehensif. Evaluasi didasarkan pada dua metrik utama: akurasi parameter (bias) dan kesesuaian distribusi statistik.

#### 4.4.1. Analisis Bias Komparatif



**Gambar 4.15** Perbandingan Bias Mean dan Standar Deviasi dari Ketiga Metode

Analisis ini berfokus pada seberapa dekat parameter data simulasi ( $n = 10.000$ ) terhadap parameter data asli. Gambar 4.15 menyajikan visualisasi perbandingan *Absolute Bias* untuk Mean (batang biru) dan Standar Deviasi (batang merah) antar metode. Berdasarkan Gambar 4.15, analisis mendalam dapat ditarik dimulai dari kinerja estimasi rata-rata (*Mean Bias*) yang ditunjukkan oleh batang biru. Metode *Acceptance-Rejection* menunjukkan kinerja terlemah dalam mengestimasi rata-rata, ditandai dengan batang biru yang paling tinggi dengan bias melebihi 1.100. Metode *Inverse-Transform* berada di posisi kedua dengan bias yang tergolong menengah. Sebaliknya, metode *Transformation (Box-Muller)* tampil sebagai yang terbaik dengan batang biru terendah (bias  $\approx 647$ ), yang berarti rata-rata data simulasinya paling mendekati rata-rata data asli BPS.

Selanjutnya, perbedaan kinerja terlihat sangat mencolok pada metrik estimasi variabilitas yang direpresentasikan oleh batang merah. Metode *Inverse-Transform* tercatat memiliki bias standar deviasi tertinggi, di mana batang merahnya hampir menyentuh angka 900. Metode *Acceptance-Rejection* menunjukkan hasil yang lebih baik, namun masih memiliki bias yang cukup terlihat di kisaran 645. Dominasi ditunjukkan oleh metode *Transformation (Box-Muller)* yang menunjukkan superioritas mutlak dengan batang merah yang sangat pendek, merefleksikan bias hanya sebesar 38,98. Secara visual dan numerik, metode Box-Muller terbukti paling stabil dan akurat dalam menjaga karakteristik statistik data dibandingkan dua metode lainnya.

#### 4.4.2. Tinjauan Goodness-of-Fit

Selain bias parameter, evaluasi juga meninjau hasil uji *Kolmogorov-Smirnov* yang telah dilakukan sebelumnya untuk melihat kesamaan bentuk distribusi secara keseluruhan. Dalam pengujian ini, metode *Acceptance-Rejection* mencatatkan p-value tertinggi sebesar 0,7121. Hal ini mengindikasikan bahwa meskipun bias parameternya sedikit besar, metode ini menghasilkan bentuk kurva distribusi yang secara geometris paling "menempel" dengan histogram data asli. Posisi berikutnya ditempati oleh metode *Inverse-Transform* dengan p-value sebesar 0,4973, diikuti oleh Transformation Method dengan p-value sebesar 0,4290. Meskipun *Transformation Method* memiliki p-value terendah, nilainya masih jauh di atas taraf signifikansi 0,05 ( $0,4290 > 0,05$ ). Artinya, secara statistik metode ini tetap valid dan distribusi yang dihasilkan terbukti sama dengan distribusi data asli.

#### 4.4.3. Pemilihan Metode Terbaik

Berdasarkan pertimbangan ganda antara akurasi parameter dan validitas distribusi, penelitian ini menyimpulkan bahwa *Transformation Method (Box-Muller)* menempati peringkat pertama dan direkomendasikan sebagai metode terbaik untuk kasus data penumpang angkutan udara Sumatera Selatan. Alasannya adalah tingkat akurasi parameter yang luar biasa tinggi, terutama pada Standar Deviasi dengan bias mendekati nol. Dalam konteks simulasi transportasi, menjaga ketepatan variabilitas (fluktuasi) sangat krusial, dan Box-Muller unggul telak di aspek ini.

Di peringkat kedua terdapat metode *Acceptance-Rejection* yang unggul dalam hal kemiripan bentuk distribusi (*p-value* tertinggi), namun memiliki kelemahan pada akurasi parameter rata-rata serta biaya komputasi yang lebih mahal akibat adanya proses penolakan data. Terakhir, metode *Inverse-Transform* menempati peringkat ketiga karena meskipun valid digunakan, kinerjanya bersifat moderat dengan bias SD paling tinggi dan tidak memiliki keunggulan spesifik dibandingkan dua metode lainnya dalam studi kasus ini. Dengan demikian, metode Box-Muller dipilih sebagai pendekatan yang paling *robust* dan efisien untuk membangkitkan data simulasi pada dataset ini.

## BAB V

### KESIMPULAN

#### 5.1. Kesimpulan

Berdasarkan hasil eksplorasi data, estimasi parameter, dan simulasi pembangkitan bilangan acak yang telah dilakukan terhadap data penumpang angkutan udara Sumatera Selatan (2015–2025), dapat ditarik beberapa kesimpulan sebagai berikut:

1. Data jumlah penumpang memiliki pola fluktuasi yang tinggi dengan komponen musiman yang kuat serta mengalami *structural break* (penurunan drastis) akibat pandemi COVID-19. Meskipun terdapat kecenderungan menjulur (*skewness*) akibat nilai ekstrim rendah saat pandemi, hasil uji statistik membuktikan bahwa data tersebut masih memenuhi asumsi distribusi normal, sehingga simulasi berbasis distribusi Normal valid untuk dilakukan.
2. Ketiga metode simulasi yang diterapkan, yaitu *Inverse-Transform*, *Acceptance-Rejection* (dengan proposal Cauchy), dan *Transformation Method (Box-Muller)*, terbukti berhasil membangkitkan data tiruan yang valid. Hal ini dikonfirmasi oleh uji *Kolmogorov-Smirnov* dimana ketiga metode menghasilkan  $p\text{-value} > 0,05$ , yang berarti tidak ada perbedaan signifikan antara distribusi data simulasi dengan data asli.
3. Berdasarkan evaluasi akurasi parameter (bias), metode *Transformation (Box-Muller)* disimpulkan sebagai metode terbaik dan paling *robust*. Metode ini menghasilkan bias estimasi rata-rata terendah dan bias standar deviasi yang nyaris sempurna (hanya 38,98 pada  $n = 10.000$ ). Hal ini menunjukkan keunggulan metode *Box-Muller* dalam mereplikasi variabilitas data secara presisi dibandingkan metode *Acceptance-Rejection* maupun *Inverse-Transform*.

#### 5.2. Saran

Berdasarkan keterbatasan yang ditemukan selama proses pengerjaan tugas ini, penulis memberikan beberapa saran untuk pengembangan atau penelitian selanjutnya:

1. Mengingat data asli memiliki sedikit kemiringan (*skewness*) akibat guncangan pandemi, penelitian selanjutnya disarankan untuk mencoba memodelkan data menggunakan distribusi yang dapat mengakomodasi *skewness* dengan lebih baik, seperti distribusi Gamma atau Weibull, untuk melihat apakah akurasi simulasi dapat lebih ditingkatkan.
2. Untuk mendapatkan parameter yang lebih representatif, disarankan untuk memisahkan analisis menjadi dua periode berbeda, yaitu periode pra-pandemi (normal) dan periode pandemi/pemulihan. Hal ini bertujuan agar varians ekstrem saat pandemi tidak mendistorsi parameter simulasi untuk kondisi normal.
3. Disarankan untuk menguji metode pembangkitan bilangan acak lain yang lebih kompleks, seperti algoritma *Marsaglia-Polar* (modifikasi *Box-Muller*) atau metode MCMC (*Markov Chain Monte Carlo*), untuk membandingkan efisiensi waktu komputasi, terutama jika bekerja dengan ukuran sampel yang sangat besar ( $n > 10^6$ ).

## DAFTAR PUSTAKA

- [1] M. N. Nasution, *Manajemen Transportasi*, Jakarta: Ghalia Indonesia, 2019.
- [2] Badan Pusat Statistik Provinsi Sumatera Selatan, "Penumpang Angkutan Udara yang Datang dan Berangkat," Palembang, BPS Sumatera Selatan, 2025.
- [3] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed., Melbourne: OTexts, 2018.
- [4] R. Horonjeff, F. X. McKelvey, W. J. Sproule and S. B. Young, *Planning and Design of Airports*, 5th ed., New York: McGraw-Hill Education, 2010.
- [5] A. M. Law, *Simulation Modeling and Analysis*, 5th ed., New York: McGraw-Hill Education, 2015.
- [6] S. M. Ross, *Simulation*, 5th ed., San Diego: Academic Press, 2013.
- [7] G. E. P. Box and M. E. Muller, "A Note on the Generation of Random Normal Deviates," *The Annals of Mathematical Statistics*, vol. 29, no. 2, p. 610–611, 1958.
- [8] X. Li, H. Zhang and X. Zhang, "Predicting model for air transport demand under uncertainties based on multiple probability distributions," *Sustainability*, vol. 14, no. 24, p. 16694, December 2022.
- [9] L. J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, 2nd ed., Belmont, CA: Duxbury Press, 1992.
- [10] Sugito, A. P. Maksum and D. Ispriyanti, "Analisis sistem antrian pesawat terbang di Bandara Internasional Ahmad Yani Semarang," *Jurnal Gaussian*, vol. 2, no. 1, pp. 21-30, 2013.
- [11] Team, R Core, "R: A Language and Environment for Statistical Computing," 2024. [Online]. Available: <https://www.R-project.org>.
- [12] M. Matsumoto and T. Nishimura, "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, p. 3–30, January 1998.
- [13] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed., New York: Springer, 2004.
- [14] J. P. C. Kleijnen, "Validation of simulation models using statistical techniques," *Journal of the Operational Research Society*, vol. 46, no. 8, pp. 1059-1065, August 1995.
- [15] F. J. Massey Jr., "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68-78, March 1951.

