

## PEMODELAN DAN SIMULASI PREDIKSI BESAR KLAIM JAMINAN KECELAKAAN KERJA DI BPJS KETENAGAKERJAAN CABANG JUANDA

**Tharisma Nova Agly Antalia**

Program Studi Matematika, FMIPA, Universitas Negeri Surabaya, Surabaya, dan Indonesia

e-mail: tharismanovaa@gmail.com

### Abstrak

BPJS Ketenagakerjaan merupakan lembaga yang memberikan perlindungan jaminan sosial bagi tenaga kerja di Indonesia, salah satunya melalui program Jaminan Kecelakaan Kerja (JKK). Setiap tahunnya, besar klaim JKK mengalami fluktuasi sehingga diperlukan prediksi untuk memperkirakan kebutuhan dana agar pengelolaan anggaran lebih optimal. Penelitian ini bertujuan memodelkan dan memprediksi besar klaim JKK pada BPJS Ketenagakerjaan Cabang Juanda menggunakan pendekatan distribusi probabilitas. Data yang digunakan berupa rekapitulasi besar klaim JKK periode Januari 2020 hingga Desember 2023, yang dibagi menjadi data latih (75%) dan data uji (25%). Distribusi Gamma, Lognormal, dan Pareto diterapkan pada data latih, dengan estimasi parameter menggunakan metode *Maximum Likelihood Estimation* (MLE). Uji kecocokan model dilakukan dengan Kolmogorov-Smirnov (KS), sedangkan pemilihan model terbaik ditentukan berdasarkan *Akaike Information Criterion* (AIC). Model terbaik kemudian digunakan untuk memprediksi besar klaim tahun 2023 dengan variasi jumlah simulasi sebanyak 50, 100, 1000, dan 10000 kali. Akurasi prediksi dievaluasi menggunakan *Mean Absolute Percentage Error* (MAPE). Hasil penelitian menunjukkan bahwa distribusi Lognormal menjadi model paling sesuai dalam memodelkan besar klaim JKK, dengan nilai AIC sebesar 1568,9080 dan nilai statistik KS sebesar 0,0756 disertai p-value 0,9763. Prediksi besar klaim tahun 2023 menggunakan model ini menghasilkan nilai MAPE terbaik sebesar 28,2518% saat jumlah simulasi sebanyak 100 kali, yang termasuk dalam kategori cukup baik. Berdasarkan hasil tersebut, distribusi Lognormal menjadi model paling tepat untuk prediksi besar klaim JKK di BPJS Ketenagakerjaan Cabang Juanda serta dapat menjadi bahan pertimbangan dalam perencanaan dana program JKK di masa mendatang.

**Kata Kunci:** Jaminan Kecelakaan Kerja, Distribusi Probabilitas, Prediksi Klaim, MLE, Uji Kolmogorov-Smirnov.

### Abstract

BPJS Ketenagakerjaan is an institution that provides social security protection for workers in Indonesia, one of which is through the Work Accident Insurance (JKK) program. Each year, the amount of JKK claims fluctuates, making it necessary to conduct predictions to estimate funding needs and optimize budget management. This study aims to model and predict the amount of JKK claims at BPJS Ketenagakerjaan Cabang Juanda using a probability distribution approach. The data used consists of a recapitulation of JKK claim amounts from January 2020 to December 2023, which were divided into training data (75%) and testing data (25%). Gamma, Lognormal, and Pareto distributions were applied to the training data, with parameter estimation carried out using the *Maximum Likelihood Estimation* (MLE) method. Model goodness-of-fit was tested using the Kolmogorov-Smirnov (KS) test, while the best model was selected based on the smallest value of the *Akaike Information Criterion* (AIC). The selected model was then used to predict claim amounts for 2023 with simulation variations of 50, 100, 1000, and 10000 iterations. Prediction accuracy was evaluated using the *Mean Absolute Percentage Error* (MAPE). The results showed that the Lognormal distribution was the most appropriate model for representing the amount of JKK claims, with an AIC value of 1568,9080 and a KS statistic of 0,0756 accompanied by a p-value of 0,9763. Predictions of the 2023 claim amounts using this model produced the best MAPE value of 28,2518% at 100 simulation iterations, which falls into the fairly good category. Based on these results, the Lognormal distribution is considered the most suitable model for predicting the amount of JKK claims at BPJS Ketenagakerjaan Juanda Branch and can serve as a reference in future planning for the JKK program's funding allocation.

**Keywords:** Work Accident Insurance, Probability Distribution, Claim Amount, MLE, Kolmogorov-Smirnov Test.

### PENDAHULUAN

Pembangunan nasional bertujuan untuk membangun masyarakat Indonesia yang sejahtera, adil, dan makmur, baik secara materiil maupun

spiritual sesuai dengan Pancasila dan Undang-Undang Dasar Negara Republik Indonesia Tahun 1945 (Undang-Undang No.13 Tahun 2003). Tenaga kerja memiliki peran penting dalam aktivitas

ekonomi, namun dihadapkan pada berbagai risiko, termasuk kecelakaan kerja akibat meningkatnya kompleksitas industri dan teknologi. Untuk menjamin perlindungan bagi tenaga kerja, pemerintah membentuk BPJS Ketenagakerjaan sebagai penyelenggara jaminan sosial tenaga kerja (Supadmo dkk., 2023). BPJS Ketenagakerjaan menyediakan lima program jaminan sosial, yaitu Jaminan Kecelakaan Kerja, Jaminan Kematian, Jaminan Hari Tua, Jaminan Pensiun, dan Jaminan Kehilangan Pekerjaan (Purnamasari dkk., 2023). Setiap perusahaan wajib memberikan perlindungan yang lebih baik bagi tenaga kerja dengan mengikuti program jaminan sosial yang ada, salah satunya Jaminan Kecelakaan Kerja. Program Jaminan Kecelakaan Kerja bertujuan memberikan manfaat pelayanan kesehatan dan santunan uang tunai bagi tenaga kerja yang mengalami kecelakaan atau penyakit akibat lingkungan kerja (PP No. 44 Tahun 2015).

Setiap tahunnya, jumlah klaim JKK mengalami fluktuasi yang cukup signifikan. Berdasarkan data BPJS Ketenagakerjaan secara nasional, jumlah klaim JKK tercatat sebesar 182.835 kasus pada tahun 2019 dan mengalami peningkatan hingga 370.747 kasus pada tahun 2023. Peningkatan klaim ini tentu berdampak pada kebutuhan dana yang harus disediakan untuk membayar klaim yang masuk, sehingga dibutuhkan perencanaan anggaran yang tepat dan akurat agar pengelolaan dana tetap optimal. Salah satu metode yang dapat digunakan untuk mendukung perencanaan dana klaim adalah melakukan prediksi terhadap besar klaim JKK yang mungkin terjadi di masa mendatang.

Sejumlah penelitian terdahulu telah dilakukan dalam konteks pemodelan distribusi klaim. Abidin dkk. (2023) meneliti penentuan premi asuransi kesehatan berdasarkan frekuensi dan besar klaim menggunakan *Generalized Linear Models* (GLM), dengan uji kecocokan model menggunakan *Maximum Likelihood Estimation* (MLE). Distribusi Gamma dan Bernoulli dinyatakan sesuai untuk data besar dan frekuensi klaim, meskipun distribusi lain belum diuji dalam penelitian tersebut. Sementara itu, Ng'elechei dkk. (2020) melakukan pemodelan frekuensi dan besaran klaim asuransi kendaraan bermotor, dengan hasil bahwa model Pareto dan Binomial Negatif paling sesuai untuk data yang digunakan. Selanjutnya, Omari dkk. (2018)

membandingkan beberapa model distribusi untuk data frekuensi dan besar klaim, dengan hasil bahwa distribusi Lognormal menjadi pilihan terbaik untuk memodelkan besar klaim. Temuan tersebut menunjukkan bahwa pemilihan distribusi sangat dipengaruhi karakteristik data.

Berdasarkan kondisi tersebut, penelitian ini bertujuan untuk memodelkan besar klaim JKK di BPJS Ketenagakerjaan Cabang Juanda menggunakan pendekatan distribusi Gamma, Pareto, dan Lognormal, kemudian memilih model terbaik berdasarkan uji kecocokan distribusi menggunakan Kolmogorov-Smirnov (KS) dan nilai *Akaike Information Criterion* (AIC). Selanjutnya, model terpilih digunakan untuk memprediksi besar klaim JKK tahun 2023 melalui proses simulasi pembangkitan data acak sesuai parameter hasil estimasi. Hasil prediksi tersebut kemudian dievaluasi menggunakan *Mean Absolute Percentage Error* (MAPE) untuk mengukur tingkat akurasi prediksi.

Penelitian ini diharapkan dapat memberikan gambaran mengenai model distribusi yang sesuai untuk memprediksi besar klaim JKK serta dapat menjadi acuan bagi BPJS Ketenagakerjaan Cabang Juanda dalam merencanakan kebutuhan dana cadangan pembayaran klaim secara lebih optimal di masa mendatang.

## KAJIAN TEORI

### KLAIM

Klaim adalah tuntutan yang dilakukan oleh tertanggung sehubungan dengan suatu kontrak, dimana tertanggung menuntut pembayaran sebagai ganti premi yang dibayarkan jika terjadi musibah, sebagai bagian dari kontrak tertanggung dengan perusahaan asuransi (Otoritas Jasa Keuangan, 2019).

### KLAIM JAMINAN KECELAKAAN KERJA (JKK)

Klaim Jaminan Kecelakaan Kerja ini merupakan proses pengajuan permohonan untuk mendapatkan manfaat berupa uang tunai dan/atau pelayanan kesehatan yang diberikan kepada peserta saat mereka mengalami kecelakaan kerja atau penyakit yang disebabkan oleh lingkungan kerja (PP No. 44 Tahun 2015).

## DISTRIBUSI PROBABILITAS

Berdasarkan karakteristik peubah acaknya, distribusi probabilitas dibagi menjadi dua jenis, yaitu distribusi probabilitas diskrit dan distribusi probabilitas kontinu (Ruhiat, 2022). Distribusi probabilitas diskrit digunakan untuk peubah acak yang hanya dapat mengambil nilai tertentu (*countable*) dengan peluang yang dinyatakan melalui fungsi massa probabilitas (PMF), sedangkan distribusi probabilitas kontinu digunakan untuk peubah acak yang dapat mengambil nilai dalam suatu interval tak terhingga dengan peluang yang dinyatakan melalui fungsi densitas probabilitas (PDF) (Walpole dkk., 2012).

Pada penelitian ini, fokus distribusi adalah distribusi kontinu karena data yang dianalisis berupa nominal besar klaim yang bersifat kontinu dan positif. Terdapat beberapa jenis distribusi kontinu yang umum digunakan dalam analisis risiko dan aktuaria, antara lain distribusi Gamma, distribusi Pareto, dan distribusi Lognormal.

### DISTRIBUSI GAMMA

Suatu peubah acak kontinu  $X$  dikatakan berdistribusi Gamma apabila memiliki parameter  $\alpha$  dan  $\beta$  yang bernilai positif, ditulis  $X \sim \text{Gamma}(\alpha, \beta)$  (Warella dkk., 2021).

Fungsi densitas probabilitas (PDF) untuk distribusi Gamma adalah:

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

Fungsi distribusi kumulatif (CDF) untuk distribusi Gamma sebagai berikut:

$$F(x; \alpha, \beta) = \frac{\gamma(\alpha, \frac{x}{\beta})}{\Gamma(\alpha)} \quad (2)$$

Nilai *mean* dan *varians* untuk distribusi Gamma masing-masing adalah  $E[X] = \alpha\beta$  dan  $\text{Var}(X) = \alpha\beta^2$ .

### DISTRIBUSI PARETO

Distribusi Pareto merupakan distribusi campuran dari distribusi Eksponensial dan distribusi Invers Gamma yang memiliki salah satu peran penting dalam penentuan risiko asuransi (Yusuf & Mutaqin, 2021).

Fungsi densitas probabilitas (PDF) untuk distribusi Pareto adalah:

$$f(x; \alpha; \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, 0 < x < \infty \quad (2)$$

Fungsi distribusi kumulatif (CDF) untuk distribusi Pareto sebagai berikut:

$$F(x; \alpha; \theta) = 1 - \left(\frac{\theta}{x}\right)^\alpha \quad (3)$$

Nilai *mean* dan *varians* untuk distribusi Pareto masing-masing adalah  $E[X] = \frac{\alpha}{\alpha-1}$  dan  $\text{Var}(X) = \frac{\theta^2\alpha}{(\alpha-1)^2(\alpha-2)}$ .

### DISTRIBUSI LOGNORMAL

Distribusi Lognormal adalah distribusi probabilitas yang digunakan untuk menggambarkan peubah acak kontinu yang logaritmanya terdistribusi normal (Omari dkk., 2018).

Fungsi densitas probabilitas untuk distribusi Lognormal adalah:

$$f(x; \mu; \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}}, x > 0 \quad (4)$$

Fungsi distribusi kumulatif (CDF) untuk distribusi Lognormal adalah:

$$F(x; \mu; \sigma) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right) \quad (5)$$

Nilai *mean* dan *varians* untuk distribusi Pareto masing-masing adalah  $E[X] = e^{\mu + \frac{\sigma^2}{2}}$  dan  $\text{Var}(X) = (e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2}$ .

### ESTIMASI PARAMETER MENGGUNAKAN MLE

Jika peubah acak  $X_1, X_2$ , dan  $X_n$  yang saling bebas dan tersebar secara identik dengan fungsi *Probability Density Function* (PDF) adalah  $f(x; \theta)$ , maka fungsi *likelihood*-nya adalah  $f(x_1; \theta), f(x_2; \theta), \dots, f(x_n; \theta)$  (Widodo dkk., 2023).

$$L(\theta; x) = \prod_{i=1}^n f(x_i; \theta), x = (x_1, \dots, x_n) \quad (6)$$

Fungsi *log-likelihood* dapat dinyatakan sebagai:

$$\ln L(\theta; x) = \sum_{i=1}^n \ln f(x_i; \theta) \quad (7)$$

Nilai parameter  $\theta$  dicari dengan cara menurunkan *log-likelihood* terhadap  $\theta$  dan menyatakannya dengan nol:

$$\frac{\partial}{\partial \theta} \ln L(\theta; x) = 0 \quad (8)$$

### UJI KESESUAIAN MODEL

Uji kesesuaian model dilakukan menggunakan Uji Kolmogorov-Smirnov (KS). Uji ini membandingkan data dan model distribusi kontinu tertentu (Yusuf & Mutaqin, 2021).

Berikut adalah hipotesis uji KS:

$H_0$ : Data berasal dari suatu populasi distribusi yang diuji

$H_1$ : Data tidak berasal dari suatu populasi distribusi yang diuji

Adapun statistik uji yang digunakan sebagai berikut (Yusuf & Mutaqin, 2021):

$$D_{KS} = \max_{1 \leq i \leq n} |F_n(x_i) - F^*(x_i)| \quad (9)$$

dengan:

- $F_n(x_i)$ : Fungsi distribusi kumulatif empiris pada data ke- $i$ , dirumuskan oleh Mardiyah dkk. (2022):

$$F_n(x_i) = \frac{1}{n} \sum_{i=1}^n (X_i \leq x) \quad (10)$$

- $F^*(x_i)$ : Fungsi distribusi kumulatif teoritis dari distribusi yang sedang diuji pada data ke- $i$ .

Keputusan uji ditentukan dengan membandingkan nilai  $D_{KS}$  dengan nilai kritis pada tingkat signifikansi tertentu. Nilai kritis bergantung pada ukuran sampel ( $n$ ) dan taraf signifikansi ( $\alpha$ ) yang digunakan, dengan pendekatan sebagai berikut (Yusuf & Mutaqin, 2021):

Tabel 1. Nilai Kritis Kolmogorov-Smirnov

Tingkat Signifikansi ( $\alpha$ )	0,10	0,05	0,01
Nilai Kritis	$\frac{1,22}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$	$\frac{1,63}{\sqrt{n}}$

$H_0$  diterima apabila nilai statistik uji  $D_{KS}$  lebih kecil dari nilai kritis pada taraf nyata  $\alpha$  yang telah ditentukan.

#### PEMILIHAN MODEL TERBAIK

*Akaike Information Criterion* (AIC) adalah metode statistik yang digunakan untuk memilih model terbaik diantara beberapa kandidat model. Model terbaik yang dipilih adalah model dengan nilai AIC yang terendah (Sani dkk., 2024).

Rumus AIC:

$$AIC = -2 \cdot \ln(L) + 2p \quad (11)$$

dengan:

$L$  : nilai maksimum fungsi *likelihood* model yang diusulkan

$p$  : jumlah parameter dalam model

#### TINGKAT KEAKURATAN PREDIKSI

Terdapat berbagai cara yang digunakan untuk menghitung tingkat kesalahan prediksi dalam model statistika, salah satunya yaitu *Mean Absolute Percentage Error* (MAPE).

MAPE merupakan pengukuran kesalahan yang menghitung persentase penyimpangan antara nilai aktual dengan nilai prediksi (Krisma dkk., 2019). Nilai MAPE dapat dihitung dengan persamaan sebagai berikut (Krisma dkk., 2019):

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right| \quad (12)$$

dengan:

$n$  : jumlah data

$X_t$  : nilai aktual pada periode ke- $t$

$F_t$  : nilai prediksi pada periode ke- $t$

Kriteria interpretasi MAPE dalam mengukur akurasi prediksi dapat dilihat pada Tabel 2 (Warmansyah & Hilpiah, 2019):

Tabel 2. Kriteria Interpretasi Nilai MAPE

Persentase MAPE	Kriteria Akurasi Prediksi
< 10%	Sangat Baik
10% – 20%	Baik
20% – 50%	Cukup
> 50%	Akurat

Nilai MAPE yang semakin kecil menunjukkan bahwa hasil prediksi semakin akurat (Warmansyah & Hilpiah, 2019).

## METODE

### JENIS PENELITIAN

Penelitian ini berjenis penelitian kuantitatif dengan pendekatan studi kasus. Pendekatan studi kasus digunakan untuk mengkaji dan menganalisis data secara menyeluruh guna membangun model distribusi dan melakukan prediksi besar klaim Jaminan Kecelakaan Kerja (JKK) di BPJS Ketenagakerjaan Cabang Juanda.

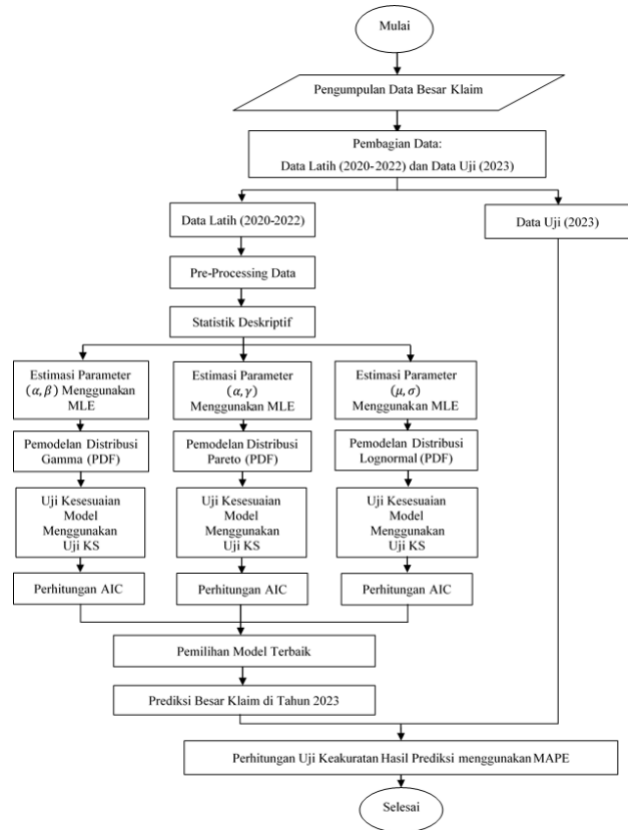
### SUMBER DATA DAN DATA PENELITIAN

Sumber data dalam penelitian ini adalah data sekunder berupa rekapitulasi besar klaim JKK bulanan yang diperoleh dari BPJS Ketenagakerjaan Cabang Juanda. Data yang digunakan merupakan data historis besar klaim JKK pada periode Januari 2020 hingga Desember 2023. Data tersebut terdiri dari dua bagian, yaitu data latih periode Januari 2020–Desember 2022 yang digunakan untuk membangun model distribusi, dan data uji periode Januari–Desember 2023 yang digunakan untuk menguji akurasi hasil prediksi.

## TEKNIK ANALISIS DATA

Adapun tahapan analisis data dalam penelitian ini adalah sebagai berikut:

1. *Pre-processing data* berupa pembersihan, imputasi data hilang, transformasi, dan normalisasi;
2. Statistik deskriptif untuk mengetahui karakteristik data melalui nilai *mean*, median, standar deviasi, dan visualisasi;
3. Estimasi parameter distribusi Gamma, Pareto, dan Lognormal menggunakan *Maximum Likelihood Estimation* (MLE);
4. Penerapan model distribusi probabilitas sesuai parameter hasil estimasi untuk membentuk model besar klaim;
5. Uji kesesuaian model menggunakan Kolmogorov-Smirnov (KS);
6. Pemilihan model terbaik berdasarkan nilai *Akaike Information Criterion* (AIC) terkecil;
7. Prediksi besar klaim JKK periode Januari-Desember 2023 menggunakan simulasi nilai acak dari model terpilih; dan
8. Pengukuran akurasi prediksi menggunakan *Mean Absolute Percentage Error* (MAPE) dengan berbagai jumlah simulasi (50, 100, 1000, dan 10000 kali).



Gambar 1. Diagram Alir Analisis Data

## HASIL DAN PEMBAHASAN

### STATISTIK DESKRIPTIF

Data yang digunakan merupakan data bulanan besar klaim Jaminan Kecelakaan Kerja (JKK) di BPJS Ketenagakerjaan Cabang Juanda periode Januari 2020-Desember 2023, terdiri dari 36 data latih dan 12 data uji. Hasil statistik deskriptif data latih disajikan pada Tabel 1.

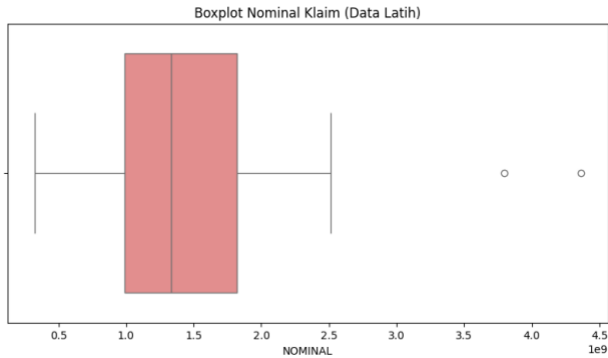
Tabel 3. Statistik Deskriptif Data Latih (Januari 2020 - Desember 2022)

Statistik Deskriptif	Nilai (Rp)
Mean	1.515.928.000
Standar Deviasi	820.878.700
Minimum	325.193.500
Kuartil 1 (Q1)	990.350.100
Median (Q2)	1.335.567.000
Kuartil 3 (Q3)	1.818.608.000
Maksimum	4.360.658.000

Hasil analisis menunjukkan bahwa nilai rata-rata klaim lebih besar daripada median, menandakan distribusi data condong ke kanan.

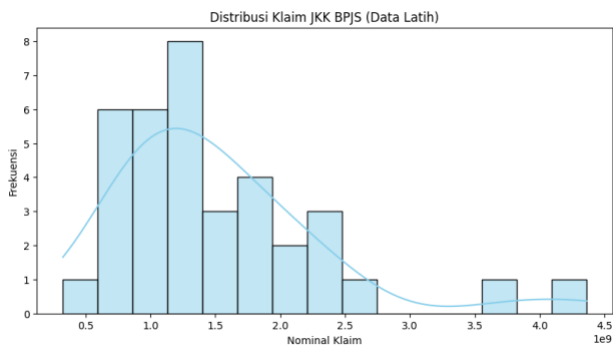
## VISUALISASI DATA

Untuk memperjelas persebaran data, visualisasi berupa *boxplot* dan histogram dengan kurva *Kernel Density Estimation* (KDE) disajikan berikut.



Gambar 2. Boxplot Besar Klaim JKK (Januari 2020–Desember 2022)

Gambar 2 menampilkan *boxplot* distribusi besar klaim JKK periode Januari 2020 hingga Desember 2022. Visualisasi ini menunjukkan bahwa sebagian besar data berada di rentang antara Rp 990 juta hingga Rp 1,8 miliar, dengan median sekitar Rp 1,3 miliar. Terdapat sejumlah klaim dengan nilai ekstrem di atas Rp 3,5 miliar yang ditandai sebagai *outlier*. Pola ini mengindikasikan distribusi data bersifat tidak simetris dan condong ke kanan (*positive skewness*).



Gambar 3. Histogram Besar Klaim JKK dengan KDE (Januari 2020–Desember 2022)

Gambar 3 memperkuat temuan tersebut melalui histogram yang dilengkapi kurva *Kernel Density Estimation* (KDE). Distribusi data terlihat memiliki ekor memanjang ke kanan, dengan konsentrasi klaim terbanyak pada interval Rp 1 miliar hingga Rp 2 miliar, sementara klaim dengan nilai sangat tinggi hanya sedikit jumlahnya.

Berdasarkan karakteristik ini, data besar klaim JKK lebih tepat dimodelkan menggunakan distribusi probabilitas yang menangani *positive skewness*, seperti Gamma, Pareto, atau Lognormal.

## ESTIMASI PARAMETER

Tiga distribusi probabilitas yaitu Gamma, Pareto, dan Lognormal diterapkan untuk memodelkan data besar klaim. Estimasi parameter dilakukan menggunakan *Maximum Likelihood Estimation* (MLE). Hasil estimasi parameter disajikan dalam Tabel 4.

Tabel 4. Hasil Estimasi Parameter Distribusi

Distribusi	Parameter	Nilai
Gamma	$\alpha$ (shape)	2,2458
	$\beta$ (scale)	757.964.080,14
Pareto	$\alpha$ (shape)	2,0000
	$\theta$ (scale)	325.193.455,00
Lognormal	$\mu$ (mean log)	21,0163
	$\sigma$ (std log)	0,4964

Hasil estimasi menunjukkan bahwa distribusi Gamma memiliki parameter  $\alpha$  sebesar 2,2458 dan  $\beta$  sebesar 757,96 juta. Distribusi Pareto memiliki nilai  $\alpha$  sebesar 2,0000 dan  $\theta$  sebesar 325,19 juta. Sementara itu, distribusi Lognormal memiliki nilai  $\mu$  sebesar 21,0163 dan  $\sigma$  sebesar 0,4964. Nilai parameter ini menggambarkan karakteristik sebaran masing-masing distribusi dan menjadi dasar dalam membentuk model prediksi besar klaim JKK.

## PENERAPAN MODEL DISTRIBUSI PROBABILITAS

Parameter hasil estimasi kemudian digunakan untuk membentuk model distribusi besar klaim JKK di BPJS Ketenagakerjaan Cabang Juanda. Model ini selanjutnya berperan sebagai dasar dalam prediksi besar klaim periode Januari–Desember 2023.

Berdasarkan hasil perhitungan, nilai parameter dari distribusi Gamma, Pareto, dan Lognormal yang telah diperoleh disubstitusikan ke dalam fungsi densitas probabilitas (PDF) masing-masing, sehingga diperoleh model distribusi berikut:

### a. Distribusi Gamma

Dengan parameter  $\alpha = 2,2458$  dan  $\beta = Rp\ 757.964.080,14$ , model distribusi Gamma dinyatakan sebagai:

$$f(x; \alpha, \beta) = \frac{1}{\Gamma(2,2458)\beta^\alpha} x^{2,2458-1} e^{-\frac{x}{757964080,14}}$$

### b. Distribusi Pareto

Dengan parameter  $\alpha = 2,0000$  dan  $\theta = 325.193.455,00$ , model distribusi Pareto dinyatakan sebagai:

$$f(x; \alpha, \theta) = \frac{2,0000 \cdot 325193455,00^{2,0000}}{x^{3,000}}$$

### c. Distribusi Lognormal

Dengan parameter  $\mu = 21,0163$  dan  $\sigma = 0,4964$ , model distribusi Lognormal dinyatakan sebagai:

$$f(x; \mu, \sigma) = \frac{1}{x \cdot 0,4964 \sqrt{2\pi}} e^{\left( -\frac{[\ln x - 21,0163]^2}{2 \cdot 0,4964^2} \right)}$$

### PENGUJIAN GOODNESS OF FIT

Uji Kolmogorov-Smirnov (KS) dilakukan untuk menilai kesesuaian model distribusi dengan data aktual. Hasil uji disajikan pada Tabel 5.

Tabel 5. Hasil Uji Kolmogorov-Smirnov

Distribusi	$D_{KS}$	$D_{krit}$	$p - value$	Keputusan
Gamma	0,1583	0,2267	0,2958	Terima
Pareto	0,7628	0,2267	0,0000	Tolak
Lognormal	0,0756	0,2267	0,9763	Terima

Distribusi Lognormal menunjukkan nilai  $p - value$  tertinggi (0,9763) dan  $D_{KS}$  terkecil, menandakan kecocokan paling baik.

### PEMILIHAN MODEL TERBAIK

Pemilihan model terbaik didasarkan pada nilai *Akaike Information Criterion* (AIC). Hasil perhitungan AIC disajikan pada Tabel 6.

Tabel 6. Hasil AIC Masing-Masing Distribusi

Distribusi	AIC
Gamma	1577,7963
Pareto	1671,2310
Lognormal	1568,9080

Distribusi dengan nilai AIC terendah dipilih sebagai model terbaik. Berdasarkan perhitungan AIC, distribusi Lognormal dengan nilai AIC 1568,9080 adalah model terbaik untuk data klaim JKK, dibandingkan dengan distribusi Gamma dan Pareto.

### PREDIKSI KLAIM DAN EVALUASI AKURASI MODEL

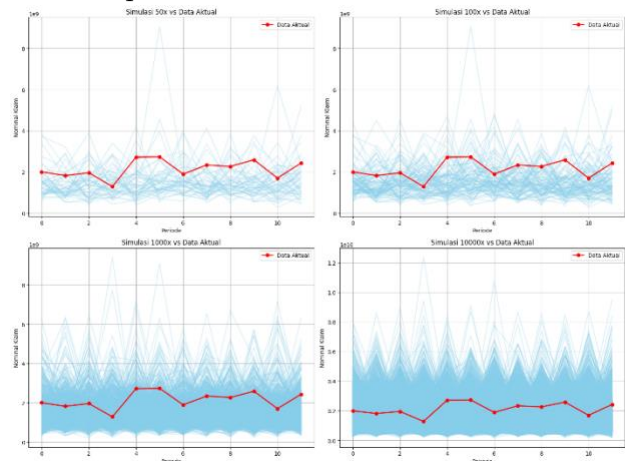
#### PROSEDUR SIMULASI PREDIKSI KLAIM

Hasil analisis distribusi menunjukkan bahwa model terbaik untuk memodelkan data besar klaim JKK di BPJS Ketenagakerjaan Cabang Juanda adalah distribusi Lognormal. Hal ini ditunjukkan oleh nilai *Akaike Information Criterion* (AIC) paling rendah sebesar 1568,9080 dibandingkan kandidat distribusi lainnya, serta hasil uji Kolmogorov-Smirnov (KS) yang menghasilkan nilai  $p - value$  sebesar 0,9763. Karena nilai  $p - value$  ini lebih besar dari tingkat signifikansi 0,05, maka dapat disimpulkan bahwa distribusi Lognormal sesuai dengan pola sebaran data besar klaim JKK yang dianalisis.

Setelah model terbaik ditetapkan, prediksi besar klaim JKK untuk periode Januari hingga Desember 2023 dilakukan dengan membangkitkan angka acak menggunakan fungsi *rvs()* dari pustaka SciPy Python, berdasarkan parameter distribusi Lognormal hasil estimasi yaitu nilai  $\mu$  dan  $\sigma$ . Untuk memastikan hasil prediksi yang stabil dan representatif, dilakukan variasi jumlah simulasi sebanyak 50, 100, 1000, dan 10000 kali pada masing-masing bulan periode prediksi.

#### SEBARAN HASIL SIMULASI

Visualisasi sebaran hasil simulasi untuk masing-masing jumlah iterasi ditampilkan pada Gambar 4. Garis merah menunjukkan nilai data aktual periode Januari hingga Desember 2023, sedangkan garis-garis biru merepresentasikan hasil simulasi.



Gambar 4. Sebaran Hasil Simulasi dan Data Aktual pada Berbagai Jumlah Simulasi

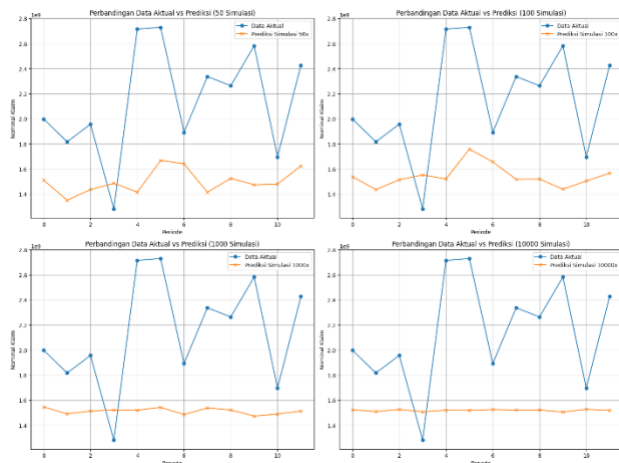
Dari gambar tersebut, terlihat bahwa seiring bertambahnya jumlah simulasi, sebaran hasil



simulasi cenderung lebih rapat di sekitar nilai tengah. Pada jumlah simulasi 50 dan 100 kali, sebaran hasilnya masih cukup lebar, sementara mulai dari 1000 hingga 10000 simulasi, pola sebaran semakin padat dengan variansi yang lebih kecil.

#### PERBANDINGAN RATA-RATA HASIL SIMULASI

Setelah proses simulasi prediksi dilakukan dengan variasi jumlah simulasi sebanyak 50, 100, 1000, dan 10000 kali, nilai hasil prediksi dari masing-masing jumlah simulasi dirata-ratakan untuk setiap bulan prediksi. Proses ini bertujuan untuk memperoleh estimasi besar klaim yang lebih stabil, konsisten, dan representatif dari keseluruhan hasil simulasi. Visualisasi perbandingan antara nilai rata-rata hasil prediksi dan data aktual untuk periode Januari hingga Desember 2023 ditampilkan pada Gambar 5 berikut.



Gambar 5. Perbandingan Data Aktual dengan Rata-rata Hasil Prediksi Simulasi

Pada grafik tersebut, garis berwarna biru merepresentasikan data aktual besar klaim JKK yang terjadi selama tahun 2023, sedangkan garis berwarna oranye menunjukkan nilai prediksi hasil rata-rata dari masing-masing jumlah simulasi.

Perbedaan nilai antar jumlah simulasi relatif kecil, meskipun prediksi dengan 100 kali simulasi menunjukkan tingkat kedekatan paling baik terhadap data aktual.

#### EVALUASI AKURASI MODEL PREDIKSI

Akurasi hasil prediksi dievaluasi menggunakan *Mean Absolute Percentage Error* (MAPE). Hasil perhitungan MAPE dari masing-masing variasi jumlah simulasi ditampilkan pada Tabel 7.

. Tabel 7. Hasil MAPE dari Berbagai Jumlah Simulasi

Jumlah Simulasi	Nilai MAPE
50 Simulasi	29,4840%
100 Simulasi	28,2518%
1000 Simulasi	29,2442%
10000 Simulasi	28,7521%

Berdasarkan Tabel 7, nilai MAPE terkecil diperoleh pada jumlah simulasi sebanyak 100 kali, yaitu 28,2518%, yang menunjukkan tingkat kesalahan persentase terkecil dibandingkan jumlah simulasi lainnya.

#### PERBANDINGAN HASIL PREDIKSI TERBAIK DENGAN DATA AKTUAL

Hasil prediksi dari 100 kali simulasi yang memiliki MAPE terkecil kemudian dibandingkan dengan data aktual. Perbandingan nilai besar klaim aktual dan hasil prediksi ditunjukkan pada Tabel 8.

Tabel 8. Perbandingan Besar Klaim JKK Aktual dan Hasil Prediksi Terbaik

Bulan	Data Aktual (Rp)	Prediksi (Rp)
Jan-23	1.997.044.520	1.536.785.480
Feb-23	1.816.157.980	1.435.291.630
Mar-23	1.956.263.880	1.512.540.470
Apr-23	1.280.485.660	1.551.621.690
Mei-23	2.714.789.000	1.519.746.240
Jun-23	2.729.161.860	1.757.790.250
Jul-23	1.891.444.820	1.655.655.400
Agu-23	2.336.039.820	1.516.834.900
Sep-23	2.263.448.250	1.518.672.380
Okt-23	2.581.912.130	1.438.294.080
Nov-23	1.695.717.700	1.503.250.690
Des-23	2.426.207.550	1.566.060.030

Dari tabel tersebut terlihat bahwa hasil prediksi dengan distribusi Lognormal melalui 100 kali simulasi secara umum cukup mendekati data aktual, meskipun masih terdapat beberapa selisih.

#### PENUTUP

#### SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa distribusi Lognormal merupakan model yang paling sesuai dalam memodelkan besar klaim Jaminan Kecelakaan Kerja (JKK) di BPJS Ketenagakerjaan Cabang Juanda.



Hal ini ditunjukkan oleh nilai *Akaike Information Criterion* (AIC) sebesar 1568,9080, yang merupakan nilai terendah dibandingkan distribusi Gamma dan Pareto. Selain itu, hasil uji Kolmogorov-Smirnov (KS) menunjukkan nilai statistik sebesar 0,0756 dengan p-value 0,9763, yang berarti tidak terdapat perbedaan signifikan antara distribusi Lognormal dan data aktual.

Distribusi Lognormal juga menghasilkan prediksi besar klaim JKK dengan tingkat akurasi yang cukup baik. Hasil pengujian menunjukkan nilai *Mean Absolute Percentage Error* (MAPE) terbaik sebesar 28,2518% saat simulasi dilakukan sebanyak 100 kali. Dengan demikian, distribusi Lognormal dapat digunakan sebagai model yang efektif dalam memprediksi besar klaim JKK dan menjadi dasar pertimbangan dalam perencanaan dana klaim di BPJS Ketenagakerjaan Cabang Juanda.

## SARAN

Berdasarkan hasil yang diperoleh, beberapa saran yang dapat diajukan untuk penelitian selanjutnya antara lain:

1. Mengkaji penggunaan distribusi alternatif seperti Weibull atau *Generalized Beta Distribution*, guna mengeksplorasi model yang mungkin lebih sesuai untuk pola klaim yang kompleks.
2. Menambahkan variabel eksternal dalam model, seperti faktor ekonomi, kebijakan pemerintah, atau kondisi sosial yang dapat memengaruhi besar klaim JKK di masa mendatang.
3. Menggunakan metode pengujian kecocokan tambahan seperti *Anderson-Darling Test* atau *Bayesian Information Criterion* (BIC) untuk memperkuat validitas pemilihan model distribusi.
4. Menerapkan metrik evaluasi tambahan seperti *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), serta melakukan analisis residual untuk mengidentifikasi potensi bias dan mengevaluasi performa model lebih komprehensif.

## DAFTAR PUSTAKA

- Abidin, S., Dewi Anugrawati, S., & Nurwahidah. (2023). Penentuan Premi Asuransi Kesehatan dengan Faktor Underwriting Menggunakan Generalized Linear Models. *MathVision : Jurnal Matematika*, 6(1), 17–22. <https://doi.org/10.55719/mv.v6i1.1109>
- Al-Qomar, L. M. A., Saepudin, D., & Rohmawati, A. A. (2018). Pemodelan dan simulasi kebangkrutan perusahaan asuransi dengan waktu kedatangan klaim berdistribusi pareto. *EProceedings of Engineering*, 5(1), 1826–1833.
- Alfaridzi, S. I. R., & Prabowo, A. (2023). PENENTUAN TARIF PREMI PADA ASURANSI KENDARAAN DENGAN BESAR KLAIM BERDISTRIBUSI EKSPONENSIAL DAN GAMMA. *Premium Insurance Business Journal*, 10(1), 29–41. <https://doi.org/10.35904/premium.v10i1.46>
- Amalia, A. N., Tirto, M. A., & Aswi. (2021). Perbandingan Metode Momen , Maximum Likelihood Dan Bayes. *VARIANSI*, 3(3), 115–125. <https://doi.org/10.35580/variensiunm26374>
- Amalia, B., Martha, S., & Rizki, S. W. (2019). ESTIMASI PARAMETER MODEL SURVIVAL DISTRIBUSI PARETO -GAMMA DENGAN METODE BAYESIAN SELF. *Buletin Ilmiah Math, Stat, Dan Terapannya (Bimaster)*, 8(3), 429–436.
- Handayani, S. (2017). PENGARUH PENYELESAIAN KLAIM ASURANSI TERHADAP PENCAPAIAN TARGET PENJUALAN PRODUK ASURANSI AJB BUMIPUTERA 1912 CABANG BENGKULU. *EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi Dan Bisnis*, 5(1), 79–85. <https://doi.org/10.37676/ekombis.v5i1.332>
- Hayuningtyasi, R. Y., & Sari, R. (2022). Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes. *Jurnal Teknik Komputer AMIK BSI*, 8(1), 40–44.
- Hilmi, M. R., Nurtiyasari, D., & Syahputra, A. (2022). Pemanfaatan Skewness dan Kurtosis dalam Menentukan Harga Opsi Beli Asia. *Quadratic: Journal of Innovation and Technology in Mathematics and Mathematics Education*, 2(1), 7–15. <https://doi.org/10.14421/quadratic.2022.021-02>
- Kharisma, D. E., Sasongko, L. R., & Mahatma, T. (2022). Model Kerugian Agregat Untuk Jaminan Kecelakaan Kerja Berdasarkan Simulasi. *Journal of Fundamental Mathematics and Applications (JFMA)*, 5(1), 67–84. <https://doi.org/10.14710/jfma.v5i1.14307>
- Krisma, A., Azhari, M., & Widagdo, P. P. (2019). Perbandingan Metode Double Exponential Smoothing Dan Triple Exponential Smoothing Dalam Parameter Tingkat Error Mean Absolute

- Percentage Error (MAPE) dan Means Absolute Deviation (MAD) Alviani Krisma Putut Pamilih Widagdo Kata kunci-forecasting, Double Ex. *Prosiding Seminar Nasional Ilmu Komputer Dan Teknologi Informasi*, 4(2), 81–87.
- Mardiyah, R., Somayasa, W., Budiman, H., Kabil Djafar, M., & Sahupala, R. (2022). Uji GOODNESS OF FIT DISTRIBUSI GAMMA TERBOBOTI DENGAN STATISTIK KOLMOGOROV-SMIRNOV UNTUK PARAMETER TERESTIMASI. *Jurnal Matematika Komputasi Dan Statistika*, 2(2), 92–101. <https://doi.org/10.33772/jmks.v2i2.13>
- Mintari, N. P. (2017). Pemodelan dan Simulasi Untuk Mengetahui Kebangkrutan Perusahaan Asuransi Berdasarkan Ukuran Klaim. *E-Proceeding of Engineering*, 4(2), 3025–3032.
- Mirawati. (2023). Prosedur Klaim Jaminan Hari Tua (Jht) Pada Bpjs Ketenagakerjaan. *Jurnal HEI EMA*, 2(2), 69–79.
- Nastiti, R. S., & Mutaqin, A. K. (2022). Penerapan Model Komposit Weibull-Pareto pada data Klaim Asuransi Harta Benda. *Jurnal Riset Statistika*, 2(1), 43–49. <https://doi.org/10.29313/jrs.vi.903>
- Ng'elechei, J. K., Cheruiyot Chelule, J., Orango, H. I., & Anapapa, A. O. (2020). Modeling Frequency and Severity of Insurance Claims in an Insurance Portfolio. *American Journal of Applied Mathematics and Statistics*, 8(3), 103–111. <https://doi.org/10.12691/ajams-8-3-4>
- Omari, C. O., Nyambura, S. G., & Mwangi, J. M. W. (2018). Modeling the Frequency and Severity of Auto Insurance Claims Using Statistical Distributions. *Journal of Mathematical Finance*, 08(01), 137–160. <https://doi.org/10.4236/jmf.2018.81012>
- Otoritas Jasa Keuangan. (2019). Buku 4 Perasuransian Seri Literasi Keuangan Perguruan Tinggi. In *Otoritas Jasa Keuangan*.
- PP No. 44 Tahun 2015 Tentang Penyelenggaraan Progrma Jaminan Kecelakaan Kerja Dan Jaminan Kematian, Lembaran Negara Republik Indonesia 1 (2015).
- Purnamasari, N. I., Susilowati, E., & Kirana, N. W. I. (2023). Analisis Penerapan Sistem Akuntansi Klaim Jaminan Kecelakaan Kerja Dan Jaminan Kematian Pada BPJS Ketenagakerjaan Cabang Juanda. *Journal of Economic, Bussines and Accounting (COSTING)*, 7(1), 872–883. <https://doi.org/10.31539/costing.v7i1.6938>
- Puspita Sari, N., & Warsito, B. (2016). Penerapan Teori Antrian Pada Pelayanan Teller Bank X Kantor Cabang Pembantu Puri Sentra Niaga. *Jurnal Gaussian*, 6(1), 81–90. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- Ruhiat, D. (2022). IMPLEMENTASI DISTRIBUSI PELUANG GUMBEL UNTUK ANALISIS DATA CURAH HUJAN RENCANA. *Teorema: Teori Dan Riset Matematika*, 7(1), 213. <https://doi.org/10.25157/teorema.v7i1.7137>
- Sani, R. F., Yanuar, F., & Devianto, D. (2024). ESTIMASI PARAMETER DARI DISTRIBUSI LOMAX MENGGUNAKAN METODE BAYESIAN ENTROPY LOSS FUCTION. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 5(3), 1441–1450.
- Sari, R. P., & Hayati MT, M. (2019). Analisis Jumlah Penerimaan Pajak Bumi dan Bangunan di Kecamatan Gunung Labuhan Way Kanan. *JTAM | Jurnal Teori Dan Aplikasi Matematika*, 3(1), 22–26. <https://doi.org/10.31764/jtam.v3i1.756>
- Setiyanto. (2024). ANALISIS HARGA TANAH PASAR WAJAR TERHADAP KEBERADAAN PUSAT PERDAGANGAN DAN JASA (Studi Kasus: Kelurahan Kauman Kec. Semarang Tengah, Kota Semarang). *Jurnal Geodesi Undip*, 13(2), 475–484.
- Sidiq, S., Maburur, N. S., Informatika, S. T., Teknik, F., & Tangerang, U. M. (2025). Pengembangan Model Prediksi Risiko Diabetes Menggunakan Pendekatan AdaBoost dan Teknik Oversampling. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (JIMA-ILKOM)*, 4(1), 13–23.
- Supadmo, H., Ghafara Rahman, H., Famelia Nining .A, D., Rama Sagita, O., & Diaz Rasendriya, R. (2023). PERAN BPJS KETENAGAKERJAAN DALAM PELAKSANAAN PROGRAM JAMINAN KECELAKAAN KERJA (JKK). *Borneo Law Review*, 6(2), 205–218. <https://doi.org/10.35334/bolrev.v6i2.3243>
- Ulkhaq, M. M., Pramono, S. N. W., & Halim, R. (2017). Aplikasi Seven Tools Untuk Mengurangi Cacat Produk Pada Mesin Communit Di PT. Masscom Graphy, Semarang. *Jurnal PASTI*, XI(3), 220–230.
- Undang-Undang No.13 Tahun 2003 Pasal 9, 1 (2014).
- UU RI No.40 Tahun 2004 Tentang Sistem Jaminan Sosial Nasional, Jdih BPK RI 1 (2004).
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). Probability and Statistics for Engineers and Scientists 9th Edition. In *United States of America: Pearson Education, Inc.*
- Warella, R. Y., Wattimanela, H. J., & Ilwaru, V. Y. I. (2021). SIFAT-SIFAT DAN KEJADIAN KHUSUS DISTRIBUSI GAMMA. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 15(1), 047–058. <https://doi.org/10.30598/barekengvol15iss1p047-058>
- Warmansyah, J., & Hilpiah, D. (2019). Penerapan

metode fuzzy sugeno untuk prediksi persediaan bahan baku. *Teknois : Jurnal Ilmiah Teknologi Informasi Dan Sains*, 9(2), 12-20. <https://doi.org/10.36350/jbs.v9i2.58>

Widodo, V. R., Adiyansyah, F., Anwar, Y. R., & Sari, K. N. (2023). Pemodelan Besar Klaim menggunakan Distribusi Berekor dan Tail-Value-at-Risk (TVaR) pada Data Sampel Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan. *Jurnal Statistika Dan Aplikasinya*, 7(2), 141-148. <https://doi.org/10.21009/JSA.07203>

Yusuf, D., & Mutaqin, A. K. (2021). Pemodelan Distribusi Pareto untuk Data Besar Klaim Asuransi Kendaraan Bermotor di Indonesia. 7, 750-755. <http://dx.doi.org/10.29313/.v0i0.30788>