

**PETA KENDALI ZERO INFLATED GENERALIZED POISSON  
MODIFIKASI EXPONENTIALLY WEIGHTED MOVING AVERAGE  
PADA DATA YANG OVERDISPERSI DAN ZERO INFLATION**

*ZERO INFLATED GENERALIZED POISSON MODIFIED  
EXPONENTIALLY WEIGHTED MOVING AVERAGE CONTROL  
CHART ON OVERDISPERSION DAN ZERO INFLATION DATA*

**RATMILA**



**PROGRAM STUDI STATISTIKA  
SEKOLAH PASCASARJANA  
UNIVERSITAS HASANUDDIN  
MAKASSAR**

**2023**

**PETA KENDALI ZERO INFLATED GENERALIZED POISSON MODIFIKASI  
EXPONENTIALLY WEIGHTED MOVING AVERAGE PADA DATA YANG  
OVERDISPERSI DAN ZERO INFLATION**

Tesis

sebagai salah satu syarat untuk mencapai gelar Magister

Program Studi Statistika

Disusun dan diajukan oleh :

RATMILA

H062212014

kepada

**PROGRAM MAGISTER STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS HASANUDDIN  
MAKASSAR  
2023**

**TESIS**

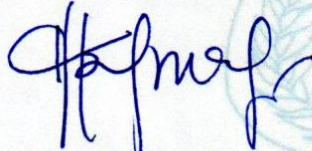
**PETA KENDALI ZERO INFLATED GENERALIZED POISSON  
MODIFIKASI EXPONENTIALLY WEIGHTED MOVING AVERAGE PADA  
DATA YANG MENGALAMI OVERDISPERSI DAN ZERO INFLATION**

**RATMILA  
H062212014**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam  
rangka Penyelesaian Program Studi Magister Statistika Fakultas  
Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin  
pada tanggal 16 Agustus 2023  
dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui

Pembimbing Utama



**Dr. Erna Tri Herdiani, S.Si., M.Si**

NIP. 19750429 200003 2 001

Pembimbing Pendamping



**Dr. Nurtiti Sunusi, S.Si., M.Si**

NIP. 19720117 199703 2 002

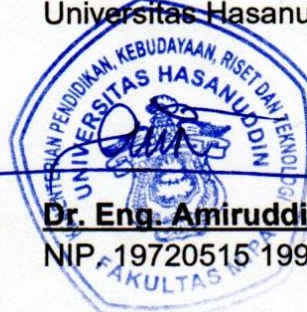
Ketua Program Studi  
Magister Statistika



**Dr. Dr. Georgina M. Tinungki, M.Si**

NIP. 19620926 198702 2 001

Dekan Fakultas Matematika dan  
Ilmu Pengetahuan Alam  
Universitas Hasanuddin



**Dr. Eng. Amiruddin, M.Si**

NIP. 19720515 199702 1 002

## PERNYATAAN KEASLIAN TESIS DAN PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa, tesis berjudul “Peta Kendali Zero Inflated Generalized Poisson Modifikasi Exponentially Weighted Moving Average Pada Data Yang Mengalami Overdispersi Dan Zero Inflation pada Data yang Overdispersi dan Zero Inflation” adalah benar karya saya dengan arahan dari komisi pembimbing Dr. Erna Tri Herdiani, S.Si, M.Si sebagai Pembimbing Utama dan Dr. Nurtiti Sunusi, S.Si., M.Si sebagai Pembimbing Pendamping. Karya ilmiah ini belum diajukan dan tidak sedang diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka tesis ini. Sebagian dari isi tesis ini telah dipublikasikan di Jurnal International Journal of Research Publications (IRJP), ISSN: 2708-3578 Vol. 129 Issue 1, July – 2023, sebagai artikel dengan judul “*A Modified Exponentially Weighted Moving Average for Monitoring Zero Inflated Generalized Poisson Processes*”.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya berupa tesis ini kepada Universitas Hasanuddin.

Makassar, 16 Agustus 2023

Yang menyatakan,



Ratmiia

NIM H062212014

## UCAPAN TERIMA KASIH

Segala Puji Syukur kepada **Allah SWT**, Tuhan Yang Maha Esa, atas berkat, bimbingan dan kasih karunia-Nya yang dilimpahkan kepada penulis, serta shalawat dan salam semoga selalu tercurahkan kepada baginda tercinta, Nabi yang paling dimuliakan, pemimpin orang-orang bertakwa, **Muhammad SAW** yang dinanti-nantikan syafaatnya di akhirat kelak. Limpahan doa kepada keluarga serta sahabat Rasulullah SAW. *Alhamdulillah*, berkat rahmat dan karunia serta mukzizat-Nya sehingga penulis dapat menyelesaikan thesis ini sebagai salah satu syarat akademik untuk meraih gelar Magister Sains pada Program Studi Magister Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

Dengan berakhirnya perjalanan panjang ini, saya ingin menyampaikan rasa terima kasih yang tulus kepada semua pihak yang telah memberikan dukungan dan bantuan selama penulisan thesis baik moril maupun materil. Penulis menyadari sepenuhnya kehadiran penulis hingga hari ini berkat doa dan restu Ibunda terkasih, **Hj.Mammi bersama adik-adikku tersayang**, yang senantiasa mengingatkan, menguatkan, dan mendukung semua keputusan penulis. Terimakasih yang tak terbatas serta doa yang senantiasa terpanjatkan atas segala bentuk kasih dan sayang kalian kepada penulis.

Terima kasih yang setulus-tulusnya dan penghargaan yang setinggi-tingginya kepada semua pihak yang telah membantu serta mendukung penulis dalam menyelesaikan thesis ini juga penulis ucapkan kepada:

1. Bapak **Prof. Dr. Ir. Jamaluddin Jompa, M.Sc.** selaku Rektor Universitas Hasanuddin.
2. Bapak **Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya.
3. Ibu **Dr. Anna Islamiyati, S.Si., M.Si** selaku Ketua Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya, serta segenap dosen pengajar dan staff Departemen Statistika.
4. Ibu **Dr. Dr. Georgina Maria Tinungki, M.Si** selaku Ketua Program Magister Statistika sekaligus salah satu Tim Penguji telah bersedia menguji serta

- memberikan saran dan kritik yang membangun, yang telah memberikan motivasi dan bantuan kepada penulis dari awal hingga selesainya penulisan tugas akhir ini.
5. Ibu **Dr. Erna Tri Herdiani, S.Si, M.Si** selaku dosen pembimbing utama dan Ibu **Dr. Nurtiti Sunusi, S.Si., M.Si** selaku pembimbing pendamping yang telah bersedia meluangkan waktu dan pikirannya untuk membimbing dan memberi arahan dengan penuh kesabaran serta selalu memotivasi penulis selama menjadi mahasiswa hingga akhir penulisan tugas akhir ini.
  6. Ibu **Sri Astuti Thamrin, S.Si., M.Stat., Ph.D** dan Bapak **Dr. Nirwan, M.Si** selaku Tim Penguji. Terima kasih telah meluangkan waktu dan telah memberikan kritik dan saran yang membangun dalam penyempurnaan penyusunan tugas akhir ini.
  7. Keluarga terkasih **Sennawi & Hj. Nurung, Andi Sugeng & Wahyuni, S.M**, serta **Alimin** yang senantiasa menjalani peran ibu dan bapak bagi penulis yang selalu menasehati, mengingatkan, dan menyemangati penulis. Terimakasih atas segala bantuannya.
  8. Bestie Lab yang teristimewa, **Andi Isna Yunita, S.Si., M.Si.**, kak **Hedi Kuswanto, S.Si., M.Si.**, Kak **Siswanto, S.Si., M.Si.**, **Ainun Utari, S.Si, Samsir Aditya Ania, S.Si., Agung Muh. Takdir, S.Si., Alimatun Najiha, S.Si., M.Si., Asnidar, S.Si.**, dan **Idman, S.Si.** Dengan penuh rasa syukur untuk setiap cerita suka duka yang penuh warna dan segala bentuk dukungan kalian yang sangat berarti penulis ucapkan terimakasih ♡
  9. Teman-teman Mahasiswa Program Magister Statistika angkatan I sampai angkatan VI, terkhusus Kak **Mubasirotul Munawarah, S.Pd.Si, Esra Rombeallo, S.Pd.,M.Si, Samsul Arifin, S.Si.,M.Si** dan kak **Nurul Azizah Muzakkir, S.Si.,M.Si** terima kasih atas bantuan dan dukungan yang luar biasa kepada penulis.
  10. Seluruh pihak yang sedikit banyaknya telah berpartisipasi dalam penulisan tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Semoga Allah SWT memberikan balasan yang berlipat ganda kepada semuanya yang telah membantu dalam penulisan tugas akhir ini. Penulis berharap semoga tulisan ini dapat memberikan manfaat bagi semua pihak yang memerlukannya. Dengan segala kerendahan hati, penulis meminta maaf atas segala kekurangan dalam penulisan tugas akhir ini.

Makassar, 5 Agustus 2023

Penulis

**ABSTRAK**

RATMILA. *Peta Kendali Zero Inflated Generalized Poisson - Modifikasi Exponentially Weighted Moving Average Pada Data Yang Mengalami Overdispersi Dan Zero Inflation* (dibimbing oleh Erna Tri Herdiani dan Nurtiti Sunusi).

Peta kendali Exponentially Weighted Moving Average (EWMA) umumnya digunakan untuk mendeteksi proses pergeseran mean secara lebih cepat. Distribusi Zero-Inflated Generalized Poisson (ZIGP) adalah distribusi yang digunakan untuk memodelkan data dengan banyak nilai nol yang berlebihan dan mengalami overdispersi, khususnya ketika menganalisis data cacah dan data diskrit. Dalam banyak kasus, data cacah memiliki nilai variansi lebih besar daripada nilai mean yang terjadi pada kasus Poisson. Distribusi ZIGP dapat menangani kasus overdispersi dengan adanya parameter tambahan pada model distribusinya. Tujuan penelitian ini adalah untuk mengadaptasi modifikasi-EWMA untuk memantau proses ZIGP sebagai alternatif untuk menangani kasus overdispersi dalam data cacah yang mengalami zero-inflasi. Estimasi parameter distribusi ZIGP menggunakan Maximum Likelihood Estimation (MLE) dan algoritma Expectation-Maximization dengan pendekatan Newton Raphson. Untuk meningkatkan efektifitas untuk mengidentifikasi proses pergeseran rata-rata yang kecil dalam data yang overdispersi, peta kendali untuk distribusi ZIGP telah dirancang dengan menggunakan statistik EWMA yang telah dimodifikasi. Nilai Average Run Length (ARL) dengan pendekatan Rantai Markov dengan nilai pembobot yang berbeda-beda digunakan untuk mengevaluasi kinerja peta kendali yang diusulkan. Hasil estimasi parameter dengan menggunakan metode MLE adalah  $\hat{\lambda} = 2.3048$ ,  $\hat{\pi} = 0.4926$ , dan  $\hat{\omega} = 0.4019$ . Sementara hasil dari nilai ARL sebagai ukuran kinerja, peta kendali ZIGP-Modifikasi EWMA yang diusulkan lebih baik digunakan daripada peta kendali ZIGP-EWMA.

**Kata Kunci:** Overdispersi, Excess Zero, distribusi Zero Inflated Generalized Poisson, Modifikasi Exponentially Weighted Moving Average, Maximum Likelihood Estimation, Average Run Length, Rantai Markov.

## ABSTRACT

RATMILA. *Zero Inflated Generalized Poisson - Modified Exponentially Weighted Moving Average Control Chart On Overdispersion and Zero Inflation Data* (supervised by Erna Tri Herdiani dan Nurtiti Sunusi).

Exponentially Weighted Moving Average (EWMA) control charts are commonly used to detect shifts in the process mean more quickly. The Zero-Inflated Generalized Poisson (ZIGP) distribution is a valuable method of statistical analysis for modeling data with excessive and overdispersion zero exceeds, particularly when analyzing counting or discrete data. In numerous cases, count data often have excessive number of zero outcomes than are expected in the Poisson. The ZIGP distribution can compensate for this overdispersion by including additional parameters that allow for greater than mean. This aim of this research is to adopt Modified-EWMA to monitoring the ZIGP processes as an alternative to handle overdispersion in zero-inflated count data. Parameter estimation ZIGP distribution using the Maximum Likelihood Estimation (MLE) and Expectation-Maximization algorithm with Newton Raphson approach. To improve the effectiveness of identifying tiny process mean shifts in overdispersion data, a control chart for the ZIGP distribution was devised in this study utilizing the modified EWMA statistic. The Average Run Length (ARL) of a Markov Chain with varied weighting constant values was used to evaluate the performance of the proposed control chart. The parameter estimation results of the ZIGP distribution using the MLE are  $\hat{\lambda} = 2.3048$ ,  $\hat{\pi} = 0.4926$ , and  $\hat{\omega} = 0.4019$ . Meanwhile the result of ARL value as a performance measure, the suggested ZIGP-Modified EWMA control chart is more suitable than the ZIGP-EWMA control chart.

**Keyword:** Overdispersion, Excess Zero, Zero Inflated Generalized Poisson distribution, Modified Exponentially Weighted Moving Average, Maximum Likelihood Estimation, Average Run Length, Markov Chain



## DAFTAR ISI

<b>SAMPUL .....</b>	<b>1</b>
<b>TESIS .....</b>	<b>ii</b>
<b>PERNYATAAN KEASLIAN TESIS.....</b>	<b>iii</b>
<b>UCAPAN TERIMA KASIH.....</b>	<b>iv</b>
<b>ABSTRAK.....</b>	<b>vi</b>
<b>ABSTRACT.....</b>	<b>vii</b>
<b>DAFTAR ISI .....</b>	<b>viii</b>
<b>DAFTAR TABEL .....</b>	<b>xi</b>
<b>DAFTAR GAMBAR .....</b>	<b>xii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xiii</b>
<b>BAB I PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
<b>BAB II TINJAUAN PUSTAKA .....</b>	<b>5</b>
2.1 Jenis-Jenis Distribusi.....	5
2.1.1 Distribusi Poisson.....	5
2.1.2 Distribusi Zero Inflated Poisson .....	8
2.1.3 Distribusi Zero Inflated Generalized Poisson .....	8
2.2 Metode Maximum Likelihood Estimation .....	9
2.3 Algoritma Expectation Maximization (EM).....	10
2.4 Uji Asumsi Data.....	12
2.4.1 Kesesuaian Distribusi.....	12
2.4.2 Overdispersi.....	13
2.5 Pengendalian Kualitas Statistik.....	14

2.5.1 Peta Kendali Exponentially Weighted Moving Average .....	14
2.5.2 Peta Kendali Modifikasi Exponentially Weighted Moving Average .....	16
2.5.3 Average Run Length.....	17
2.6 Penambangan/Pemindahan Puncak Gunung & Salamander .....	19
2.7 Kerangka Konseptual .....	22
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>23</b>
3.1 Sumber Data .....	23
3.2 Identifikasi Variabel .....	23
3.3 Metode Analisis Data .....	23
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>	<b>25</b>
4.1 Estimasi Parameter Distribusi Zero-Inflated Generalized Poisson .....	25
4.2 Peta Kendali ZIGP EWMA .....	37
4.2.1 Penentuan Mean dan Variansi untuk $Zt$ .....	38
4.2.2 Penentuan Batas Kendali ZIGP EWMA .....	41
4.3 Peta Kendali ZIGP Modifikasi-EWMA.....	42
4.2.1 Penentuan Mean dan Variansi untuk $Mt$ .....	43
4.2.2 Penentuan Batas Kendali ZIGP Modifikasi-EWMA.....	47
4.4 Aplikasi Pada Data .....	47
4.4.1 Hasil Estimasi Paramater.....	47
4.4.2 Uji Kesesuaian Distribusi .....	47
4.4.3 Uji Overdispersitas.....	48
4.4.4 Peta Kendali ZIGP- Exponentially Weighted Moving Average .....	49
i. Menentukan Nilai $Zt$ pada Peta Kendali ZIGP-EWMA untuk $\theta = 0.01$ .....	49
ii. Menentukan Nilai $Zt$ pada Peta Kendali ZIGP-EWMA untuk $\theta = 0.02$ .....	50
4.4.5 Peta Kendali ZIGP - Modifikasi Exponentially Weighted Moving Average.....	52
i. Menentukan Nilai $Mt$ pada Peta Kendali ZIGP-Modifikasi EWMA untuk $\theta =$ 0.01.....	52
ii. Menentukan Nilai $Mt$ pada Peta Kendali ZIGP-EWMA untuk $\theta=0.02$ .....	54
4.4.6 Menentukan Nilai ARL .....	57
i. ARL ZIGP-EWMA.....	58

ii. ARL ZIGP-Modifikasi EWMA.....	59
<b>BAB V PENUTUP.....</b>	<b>63</b>
5.1 Kesimpulan.....	63
5.2 Saran.....	64
<b>DAFTAR PUSTAKA.....</b>	<b>65</b>
<b>LAMPIRAN.....</b>	<b>69</b>

**DAFTAR TABEL**

<b>Tabel 4.1</b> Hasil estimasi parameter.....	47
<b>Tabel 4.2</b> Uji Kecocokan Distribusi .....	48
<b>Tabel 4.3</b> Uji Overdispersi .....	48
<b>Tabel 4.4</b> Batas Kendali ZIGP-EWMA untuk $\theta = 0.01$ .....	49
<b>Tabel 4.5</b> Batas Kendali ZIGP-EWMA untuk $\theta = 0.02$ .....	51
<b>Tabel 4.6</b> Batas Kendali ZIGP- Modifikasi EWMA untuk $\theta = 0.01$ .....	53
<b>Tabel 4.7</b> Batas Kendali ZIGP- Modifikasi EWMA untuk $\theta = 0.02$ .....	54
<b>Tabel 4.8</b> Jumlah Out Of Control.....	55
<b>Tabel 4.9</b> Nilai ARL untuk masing-masing peta kendali .....	61

## DAFTAR GAMBAR

<b>Gambar 4.1</b>	Peta kendali ZIGP-EWMA untuk $\theta = 0.01$ .....	50
<b>Gambar 4.2</b>	Peta kendali ZIGP-EWMA untuk $\theta = 0.02$ .....	51
<b>Gambar 4.3</b>	Peta Kendali ZIGP-Modifikasi EWMA untuk $\theta = 0.01$ .....	53
<b>Gambar 4.4</b>	Peta Kendali ZIGP-Modifikasi EWMA untuk $\theta = 0.02$ .....	55
<b>Gambar 4.5</b>	Grafik jumlah pengamatan yang out of control .....	56
<b>Gambar 4.6</b>	Nilai ARL Peta Kendali ZIGP-EWMA dan ZIGP-Modifikasi EWMA	62

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Hasil Estimasi Parameter .....	70
<b>Lampiran 2.</b> Uji Kecocokan Distribusi.....	72
<b>Lampiran 3.</b> Uji Overdispersi .....	73
<b>Lampiran 4.</b> Daftar Riwayat Hidup .....	74

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Peta kendali merupakan salah satu alat Pengendalian Kualitas Statistik yang secara luas banyak digunakan untuk memantau kualitas proses produksi. Peta kendali Shewhart, grafik *Exponentially Weighted Moving Average* (EWMA), dan grafik CUSUM sering digunakan untuk mendeteksi pergeseran dalam pengamatan yang normal dan saling bebas dengan variansi umum yang berasal dari proses tertentu. Literatur tentang evaluasi proses dari ketiga bagan ini menunjukkan bahwa peta kendali Shewhart mendeteksi semua pergeseran besar, peta kendali CUSUM mendeteksi pergeseran dengan mengumpulkan secara langsung semua informasi di dalam barisan nilai nilai sampel dengan menggambarkan jumlah kumulatif deviasi sampel, sedangkan peta kendali EWMA mendeteksi pergeseran melalui perubahan yang terakumulasi di bawah pemulusan eksponensial. Secara umum, ketiga peta kendali ini gagal mendeteksi pergeseran dalam satu kasus, baik pergeseran besar maupun pergeseran kecil dalam satu proses. (Patel & Divecha, 2011).

Secara khusus, sebagian besar proses industri sensitif terhadap perubahan kecil. Sebuah perubahan secara tiba-tiba adalah perubahan besar yang tidak terduga terjadi pada proses yang dipantau untuk pergeseran kecil. EWMA banyak digunakan secara luas dalam pemodelan deret waktu dan peramalan karena EWMA dapat dilihat sebagai rata-rata terboboti dari semua pengamatan dan tidak sensitif terhadap asumsi normalitas (Box dkk, 2016).

Banyak peneliti Pengendalian Kualitas Statistik telah mengembangkan beberapa metode untuk meningkatkan deteksi kemampuan peta kendali EWMA untuk pergeseran kecil hingga sedang. Yaschin (1995) telah membahas estimasi dari suatu proses pada perubahan kecil serta perubahan mendadak atas dasar EWMA. Capizzi & Masarotto (2012) telah mengembangkan peta kendali baru dari peta kendali EWMA yang disebut peta kendali *Adaptive*-EWMA untuk mendeteksi semua jenis pergeseran pada data yang saling bebas. Reynolds & Stoumbos (2006) menyarankan menggunakan dua peta kendali EWMA secara bersamaan, satu untuk proses pergeseran mean dan satu lagi untuk menangani proses pergeseran variansi yang mendeteksi pergeseran kecil dengan perubahan mendadak. Shawky & Amal (1992) mengembangkan peta kendali *Double* EWMA (DEWMA) yang telah dipelajari dan

digunakan oleh banyak penulis, seperti Mahmoud & Woodall (2010) membandingkan peta kendali EWMA dengan DEWMA, Kho dkk (2010) mengembangkan peta kendali Max-DEWMA untuk mendeteksi pergeseran kecil dan sedang, serta Alevizakos & Koukouvinos (2021) mempelajari peta kendali Modifikasi EWMA menggunakan berbagai nilai parameter tambahannya dan juga mengusulkan peta kendali *Double Modifikasi EWMA* (DMEWMA) untuk memantau pergeseran rata-rata proses, dengan asumsi bahwa prosesnya mengikuti distribusi normal.

Patel & Divecha (2011) memperkenalkan Modifikasi EWMA yang mempertimbangkan pengamatan sebelumnya dan perilaku proses saat ini dalam pengamatan untuk mendeteksi pergeseran dini suatu proses produksi secara efektif. Fandrilla dkk. (2018) menyelidiki peta kendali Modifikasi EWMA pada data yang berautokorelasi. Borrer (1998) memperkenalkan peta kendali Poisson EWMA dengan asumsi data berdistribusi Poisson. Aslam dkk. (2017) mengembangkan peta kendali Modifikasi EWMA untuk memantau proses *Conway Maxwell-Poisson*. Sementara itu, Aly dkk (2019) memantau sebuah proses dengan data yang berdistribusi Zero-Inflated Poisson dengan menggunakan peta kendali Adaptif EWMA. Jamaluddin Afgani (2021) memantau data dengan distribusi *Zero Inflated Generalized Poisson* (ZIGP) dan mengembangkan peta kendali EWMA untuk menangani kasus overdispersi.

Overdispersi adalah keadaan jika variansi lebih besar daripada *mean*. Kasus overdispersi ditemukan pada data yang melanggar asumsi equidispersi dari distribusi Poisson. Kasus overdispersi pada awalnya dapat diatasi oleh distribusi Generalized Poisson (GP). Namun, distribusi GP dapat mengatasi masalah overdispersi tetapi tidak dapat mengatasi masalah *zero inflated* atau kasus dengan data yang banyak mengandung nilai nol. Oleh karena itu diperlukan suatu model yang dapat menangani masalah tersebut. Salah satu distribusi yang dapat menangani masalah *zero inflated* adalah distribusi *Zero Inflated Poisson* (ZIP) (Lambert, 1992). Distribusi ZIP merupakan distribusi yang dapat digunakan pada data cacah dengan frekuensi nol lebih banyak. Kusuma dkk (2013) menerapkan model regresi ZIP pada data overdispersi. Akan tetapi, model ZIP ini kurang tepat untuk mengatasi masalah *overdispersi* sehingga diperlukan suatu model alternatif lain yang tepat untuk mengatasi permasalahan tersebut. Salah satunya adalah memodelkan data cacah tersebut ke dalam model distribusi *Zero Inflated Generalized Poisson* (ZIGP). Menurut Famoye & Singh (2006) distribusi ZIGP merupakan perluasan dari distribusi Poisson dan merupakan distribusi gabungan dari distribusi ZIP dan distribusi GP. Oleh



karena itu, distribusi ZIGP dapat diterapkan pada data cacah yang menunjukkan sifat *overdispersi* serta mempunyai frekuensi nol yang lebih banyak. Siska dkk (2014) dan Sulistyaningsih dkk (2019) menerapkan regresi ZIGP untuk menangani data yang mengalami *overdispersi*.

*Maximum Likelihood Estimator* (MLE) adalah salah satu metode penaksiran parameter yang dapat digunakan untuk menaksir parameter suatu model yang diketahui distribusinya. Metode MLE dilakukan dengan cara memaksimumkan fungsi *likelihood*. Penaksiran parameter model regresi ZIP dilakukan dengan metode MLE (Jansakul & Hinde, 2002). Penaksiran parameter dengan metode penaksir MLE dihitung dengan memaksimumkan fungsi *ln-likelihood*-nya. Penjumlahan fungsi *ln-likelihood* tidak dapat diselesaikan dengan metode numerik biasa, oleh karena itu dapat digunakan algoritma EM (*Expectation Maximization*).

Kinerja peta kendali untuk menilai sensitifitasnya disebut *Average Run Length* (ARL). Lucas & Saccucci (1990) menghitung *run-length* dari grafik EWMA menggunakan pendekatan rantai Markov dan mempresentasikan program komputer untuk perhitungan rata-rata *run-length* untuk skema EWMA. Steiner dkk (1999) mempelajari sifat *run-length* dari peta kendali EWMA dengan batas kendali yang waktunya berubah-ubah menggunakan rantai Markov yang tidak homogen dan mengusulkan fitur *Fast Initial Response* (FIR) untuk mendeteksi pergeseran cepat pada awal proses. Dalam membandingkan kinerja peta kendali *Shewart* dan EWMA, dan EWMA-MA pada penelitian Areepong dkk (2020) menggunakan ARL, SDRL, dan MRL.

Berdasarkan penelitian yang dilakukan oleh Khan dkk (2016) lebih banyak merinci pengembangan peta kendali Modifikasi EWMA dan menyertakan parameter tambahan yang berkaitan dengan kesalahan antara dua pengamatan dengan asumsi bahwa data berdistribusi normal atau berautokorelasi. Atas dasar studi literatur pendahuluan yang telah dilakukan sebelumnya, peneliti tertarik untuk membuat sebuah peta kendali modifikasi EWMA dengan menggunakan data yang berdistribusi ZIGP dengan judul penelitian “**Peta Kendali Zero Inflated Generalized Poisson - Modifikasi Exponentially Weighted Moving Average pada Data yang Overdispersi dan Zero Inflation (Studi Kasus: Habitat Salamander daerah Pertambangan)**”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka didapatkan rumusan masalah sebagai berikut:

1. Bagaimana hasil estimasi parameter distribusi ZIGP dengan menggunakan metode MLE?
2. Bagaimana rumusan batas kendali ZIGP- EWMA dan ZIGP-Modifikasi EWMA?
3. Bagaimana efektifitas kinerja peta ZIGP-EWMA dan ZIGP-Modifikasi EWMA berdasarkan nilai ARL?

## 1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Metode estimasi yang digunakan adalah *Maximum Likelihood Estimation*.
2. Perhitungan ARL dilakukan dengan pendekatan *Markov Chain*.

## 1.4 Tujuan Penelitian

Tujuan yang ingin dicapai pada penelitian ini adalah:

1. Memperoleh hasil estimasi parameter distribusi ZIGP
2. Memperoleh rumusan peta kendali ZIGP-EWMA dan ZIGP-Modifikasi EWMA pada data Salamander.
3. Memperoleh nilai ARL peta kendali ZIGP-EWMA dan ZIGP-Modifikasi EWMA.

## 1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Menambah wawasan, pengetahuan, dan pemahaman teoritis maupun praktis mengenai peta kendali Modifikasi-EWMA menggunakan distribusi ZIGP pada khususnya serta dapat memotivasi dalam pengembangan pengetahuan tentang peta kendali kualitas menggunakan berbagai jenis distribusi serta metode peta kendali kualitas pada umumnya.
2. Memberi tambahan informasi yang dapat dijadikan sebagai salah satu acuan pengambilan keputusan bagi para penambang, ahli ekologi, dan ahli lingkungan untuk mempertimbangkan peraturan baru yang menyempurnakan aturan lama yang mengatur tentang regulasi, manajemen, pemanfaatan penambangan bagian atas gunung agar tetap menjaga habitat/ekosistem salamander.

## BAB II TINJAUAN PUSTAKA

### 2.1 Jenis-Jenis Distribusi

#### 2.1.1 Distribusi Poisson

Distribusi Poisson adalah distribusi probabilitas diskrit yang menyatakan peluang jumlah peristiwa yang terjadi pada periode waktu tertentu apabila rata-rata kejadian tersebut diketahui dan dalam waktu yang saling bebas sejak kejadian terakhir. Distribusi Poisson juga dapat digunakan untuk jumlah kejadian pada interval tertentu seperti jarak, luas, atau volume.

Bilangan  $X$  yang menyatakan banyaknya hasil percobaan dalam suatu percobaan Poisson disebut variabel acak Poisson dan distribusi peluangnya disebut distribusi Poisson (Walpole, 2015). Karena nilai-nilai peluangnya hanya bergantung pada  $\lambda$ , yaitu rata-rata banyaknya hasil percobaan yang terjadi selama selang waktu atau daerah yang diberikan atau nilai harapan kejadian pada suatu interval adalah  $\lambda$ , maka probabilitas terjadi peristiwa sebanyak  $x$  kali ( $x$  adalah bilangan bulat non negatif,  $x = 0,1,2, \dots$ ) maka sama dengan

$$F(X = x) = f(x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!}, & x = 0,1,2, \dots \\ 0 & , x \text{ lainnya} \end{cases} \quad (2.1)$$

dengan

$x$ : banyaknya kejadian suatu peristiwa

$X$ : variabel acak

$\lambda$ : mean

Mean dan variansi dari variabel acak  $X$  yang mengikuti distribusi Poisson dengan parameter  $\lambda$  masing-masing adalah:

$$E[X] = \lambda \text{ dan } Var(X) = \lambda$$

Untuk membuktikan  $E[X] = \lambda$  dan  $Var(X) = \lambda$  dengan fungsi pembangkit moment atau *Moment Generating Function* (MGF) dari distribusi poisson adalah sebagai berikut:

Pembahasan awal dari bagian ini adalah menurunkan persamaan MGF-nya. Selanjutnya menurunkan momen pertama dan momen kedua berdasarkan persamaan MGF yang telah diperoleh sebelumnya.

1. MGF Distribusi Poisson

$$M_x(t) = E(e^{tx})$$

$$\begin{aligned} M_x(t) &= \sum_{x=0}^n e^{tx} f(x) = \sum_{x=0}^n e^{tx} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^n \frac{(\lambda e^t)^x}{x!} \\ &= e^{-\lambda} e^{\lambda e^t} \sum_{x=0}^n \frac{e^{-\lambda e^t} (\lambda e^t)^x}{x!} \end{aligned}$$

Karena  $\sum_{x=0}^n \frac{e^{-\lambda e^t} (\lambda e^t)^x}{x!} = 1$ , maka diperoleh hasil sebagai berikut:

$$M_x(t) = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

2. Mean

$$E(X) = M'_x(t = 0)$$

$$= \frac{d}{dt} e^{\lambda(e^t-1)} |_{t=0}$$

$$= \lambda e^t e^{\lambda(e^t-1)} |_{t=0}$$

$$E(X) = \lambda e^0 e^{\lambda(e^0-1)} = \lambda$$

3. Variansi

$$Var(X) = E[(X - E(X))^2]$$

$$E(X^2) = M''_x(t = 0) = \frac{d^2}{dt^2} e^{\lambda(e^t-1)} |_{t=0}$$

$$= \frac{d}{dt} \lambda e^t e^{\lambda(e^t-1)} |_{t=0}$$

$$= (\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)} |_{t=0}$$

$$= (\lambda e^0)^2 e^{\lambda(e^0-1)} + \lambda e^0 e^{\lambda(e^0-1)} = \lambda^2 + \lambda = \lambda(\lambda + 1)$$

$$E[(X - E(X))^2] = E[X^2 - 2XE(X) + (E(X))^2]$$

$$= E(X^2) - 2E(X)E(X) + (E(X))^2$$

$$= \lambda^2 + \lambda - 2\lambda\lambda + \lambda^2 = 2\lambda^2 + \lambda - 2\lambda^2 = \lambda$$

Dari momen pertama dan kedua tersebut dapat diketahui rata-rata (mean) dan variansi dari distribusi poisson.

Dalam proses Poisson, jika peristiwa terjadi rata-rata pada laju  $\lambda$  dalam satuan waktu, maka akan ada rata-rata  $\lambda t$  kejadian pada  $t$  satuan waktu. Cooper (2005) menggambarkan proses ini dalam bentuk distribusi Poisson sebagai berikut:

$$f(x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}, \quad x = 0, 1, 2, \dots \quad (2.2)$$

dengan  $f(x = 0) = e^{-\lambda t}$  adalah peluang tidak ada kejadian dalam  $t$  satuan waktu.

Menurut Walpole (1992), beberapa karakteristik distribusi Poisson adalah sebagai berikut.

1. Banyaknya kejadian yang terjadi dalam suatu interval tertentu atau suatu daerah tertentu tidak bergantung pada kejadian yang terjadi pada interval lain atau daerah lain yang terpisah (tidak berpotongan dan independen, dalam kaitan ini, proses *Poisson* dikatakan tidak punya ingatan).
2. Peluang terjadinya suatu kejadian selama interval waktu yang singkat atau dalam suatu daerah yang kecil, sebanding dengan panjang interval waktu tersebut atau besarnya daerah tersebut dan tidak bergantung pada banyaknya kejadian yang terjadi di luar interval waktu atau daerah tersebut.
3. Peluang bahwa lebih dari satu kejadian akan terjadi dalam interval waktu yang singkat tersebut atau dalam daerah yang kecil tersebut dapat diabaikan.

Pengembangan dari distribusi Poisson pertama kali diperkenalkan oleh Consul & Famoye (1992) yang dikenal sebagai distribusi *Generalized Poisson* (GP). Distribusi GP merupakan suatu model yang digunakan jika terjadi pelanggaran asumsi pada distribusi Poisson yaitu *over/under* dispersi. Distribusi GP dinyatakan sebagai berikut:

$$f(\mu, \omega, x_i) = \begin{cases} \left(\frac{\lambda}{1 + \omega\lambda}\right)^{x_i} \frac{(1 + \omega x_i)^{x_i - 1}}{x_i!} \exp\left[\frac{-\lambda(1 + \omega x_i)}{1 + \omega\lambda}\right], & x_i = 0, 1, 2, \dots, n \\ 0 & , \quad x_i \text{ yang lain.} \end{cases} \quad (2.3)$$

Mean dan Variansi dari distribusi GP adalah sebagai berikut:

$$E[X] = \lambda$$

$$Var(X) = \lambda(1 + \omega\lambda)^2$$

dimana  $\omega$  merupakan parameter dispersi dan  $x_i = 0, 1, 2, \dots$ , merupakan variabel yang berdistribusi GP. Jika  $\omega = 0$  maka model distribusi GP akan menjadi distribusi *Poisson*. Jika  $\omega > 0$  maka distribusi GP merepresentasikan data *count* yang

overdispersi. Jika  $\omega < 0$  maka model distribusi GP merepresentasikan data *count* yang underdispersi.

### 2.1.2 Distribusi Zero Inflated Poisson

Kata “Zero-Inflated” menggambarkan situasi di mana frekuensi kemunculan angka nol lebih tinggi dibandingkan dengan kemunculan angka lainnya. Menurut Famoye & Singh (2006) proporsi kemunculan nilai nol adalah sebesar 63,7% untuk data berdistribusi *Zero-Inflated Poisson* (ZIP). Distribusi ZIP adalah distribusi sederhana untuk data diskrit dengan banyak peristiwa nol (Lambert (1992) serta Vasileios Alevizakos & Christos Koukouvinos (2019)). Fungsi probabilitas dari distribusi ZIP adalah sebagai berikut:

$$f_{ZIP}(x) = \begin{cases} \pi + (1 - \pi)e^{-\lambda} & , x_i = 0 \\ \frac{(1 - \pi)e^{-\lambda}\lambda^{x_i}}{x_i!} & , x_i > 0 \end{cases} \quad (2.4)$$

dengan

$\lambda$ : mean

$\pi$ : probabilitas data pada saat  $x_i = 0$

Bila  $X$  merupakan variabel acak dengan fungsi probabilitas ZIP, maka:

untuk  $x_i = 0$

$$E[X] = Var(X) = 0$$

untuk  $x_i > 0$

$$E[X] = \mu = (1 - \pi)\lambda,$$

$$Var(X) = (1 - \pi)[\lambda^2 + \lambda] - (1 - \pi)^2\lambda^2 = E[X](1 + \pi\lambda)$$

### 2.1.3 Distribusi Zero Inflated Generalized Poisson

Distribusi *Zero-Inflated Generalized Poisson* (ZIGP) merupakan salah satu distribusi yang dapat digunakan untuk data respon yang bersifat cacah. Famoye dan Singh (2006) mendefinisikan fungsi probabilitas ZIGP sebagai gabungan dari fungsi probabilitas ZIP dan GP. Fungsi probabilitas distribusi ZIGP dapat dituliskan sebagai berikut:

$$f_{ZIGP}(x_i) = \begin{cases} \pi + (1 - \pi) \exp\left[\frac{-\lambda}{1 + \omega\lambda}\right], & x_i = 0 \\ (1 - \pi) \left(\frac{\lambda}{1 + \omega\lambda}\right)^{x_i} \frac{(1 + \omega x_i)^{x_i - 1}}{x_i!} \exp\left[\frac{-\lambda(1 + \omega x_i)}{1 + \omega\lambda}\right], & x_i > 0 \end{cases} \quad (2.5)$$

dengan

$\lambda$  : mean

$\omega$  : parameter dispersi

$\pi$  : probabilitas data pada saat  $x_i = 0$

Mean dan variansi dari distribusi ZIGP adalah sebagai berikut:

$$E[X] = (1 - \pi) \lambda \quad (2.6)$$

$$\begin{aligned} Var(X) &= (1 - \pi)[\lambda^2 + \lambda(1 + \omega\lambda)^2] - (1 - \pi)^2 \lambda^2 \\ &= E[X][(1 + \omega\lambda)^2 + \pi\lambda] \end{aligned} \quad (2.7)$$

## 2.2 Metode Maximum Likelihood Estimation

Metode *Maximum Likelihood Estimation* (MLE) adalah salah satu metode penaksiran parameter yang dapat digunakan untuk menaksir parameter suatu model yang diketahui distribusinya. Misalkan suatu sampel acak  $X_1, X_2, \dots, X_n$  dari suatu distribusi yang memiliki fungsi probabilitas dengan satu parameter  $f(X_i; \theta)$ , maka fungsi probabilitas gabungan dari sampel acak  $X_1, X_2, \dots, X_n$  adalah:

$$f(x_1, x_2, \dots, x_n, \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta) \quad (2.13)$$

Fungsi *likelihood* didefinisikan sebagai fungsi probabilitas bersama dari  $X_1, X_2, \dots, X_n$  yang dapat dianggap sebagai fungsi  $\theta$ . Misalkan fungsi *likelihood*:

$$L(\theta; x_1, x_2, \dots, x_n) = L(\theta)$$

$$L(\theta) = f(x_1, x_2, \dots, x_n, \theta) = f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta); \theta \in \Omega$$

$$= \prod_{i=1}^n f(x_i; \theta) \quad (2.14)$$

Penaksiran maksimum *likelihood* yang memaksimumkan fungsi *likelihood*  $\hat{\theta}$  disebut estimator maksimum *likelihood* dari

$\theta$ , untuk memaksimumkan fungsi *likelihood*  $\hat{\theta}$  lebih mudah bekerja dengan logaritma natural dari fungsi *likelihood*, yaitu  $l(\theta) = \ln L(\theta)$ . Nilai  $\theta$  yang

memaksimumkan  $l(\theta)$  dapat diperoleh dengan mencari solusi dari persamaan  $\frac{\partial l(\theta)}{\partial \theta} = 0$  (Hoog & Craig, 1995).

Jika pada proses estimasi parameter didapatkan persamaan terakhir yang non-linier maka tidak mudah untuk memperoleh estimasi tersebut, sehingga diperlukan suatu algoritma dan metode iterasi untuk menyelesaikan persamaan non-linier tersebut, salah satunya dengan menggunakan algoritma Expectation Maximization dengan iterasi Newton-Raphson.

### 2.3 Algoritma Expectation Maximation (EM)

Algoritma Expectation Maximization (EM) pertama kali diperkenalkan oleh Dempster, Lairt, dan Rubin pada tahun 1977. Algoritma EM digunakan untuk mengestimasi suatu parameter dengan memaksimumkan fungsi  $\ln - likelihood$  yang mengandung data hilang atau data yang tidak lengkap (*incomplete data*). Algoritma EM terdiri dari dua tahap, yaitu tahap Ekspektasi (*E-Step*) untuk mencari nilai ekspektasi dari fungsi *likelihood* dan tahap Maksimalisasi (*M-Step*) untuk memaksimumkan fungsi yang telah didefinisikan pada tahap ekspektasi sehingga didapatkan estimator parameter yang konvergen (Kusuma, 2013). Algoritma EM juga lebih mudah diterapkan ketika masalah optimasi memiliki banyak parameter dibandingkan dengan metode iterasi lainnya.

Misalnya diasumsikan data pengamatan  $x$  yang berdistribusi tertentu yang mengandung data tidak lengkap  $z$ , untuk mengatasinya dibentuk distribusi gabungan antara  $x$  dan  $z$ , yaitu:

$$f(x, z|\theta) = f(z).f(x|z)$$

Tahap Ekspektasi (*E-Step*) pada algoritma EM dilakukan dengan menghitung ekspektasi dari fungsi  $\ln - likelihood$  data tidak lengkap berdasarkan data pengamatan yang ada, yang digunakan untuk mengganti keberadaan data yang dianggap hilang, sehingga didefinisikan sebagai berikut:

$$Q(\theta|\theta^{(0)}, x) = E[\ln L(\theta^{(0)}|x, z)]$$

Pada tahap Maksimalisasi (*M-Step*) dilakukan dengan mencari nilai estimator yang dapat memaksimumkan fungsi  $Q$  yang telah didefinisikan pada tahap ekspektasi. Metode yang dapat digunakan untuk memaksimumkan fungsi  $Q$  adalah metode iterasi *Newton-Raphson* yang dilakukan secara numerik. Menurut Famoye & Singh (2006), metode *Newton-Raphson* merupakan metode numerik untuk menyelesaikan



persamaan non-linier secara iteratif seperti menyelesaikan persamaan *likelihood* yang mencari lokasi untuk memaksimalkan suatu fungsi. Jika ada proses estimasi parameter didapatkan persamaan terakhir yang non-linier maka tidak mudah untuk memperoleh estimasi parameter tersebut, sehingga diperlukan metode Newton-Raphson untuk menyelesaikan persamaan non-linier tersebut. Dasar dari metode ini adalah pendekatan deret Taylor linier.

$$f(x) = f(x^{(0)}) + \sum_{i=1}^n \frac{\partial^i f(x^{(0)})}{i! \partial(x^{(0)})^i} (x - x^{(0)})^i$$

Perluasan dari bentuk orde-1  $\frac{\partial f(\theta)}{\partial \theta} = 0$ , diperoleh

$$\begin{aligned} \frac{\partial f(\theta)}{\partial \theta} &= f(\theta^{(0)}) + H^{(0)}(\theta - \theta^{(0)}) \\ &= G^{(0)} + H^{(0)}(\theta - \theta^{(0)}) = 0 \end{aligned}$$

Jika  $\theta^{(0)}$  merupakan nilai awal dari  $\theta$  maka dapat dimisalkan  $\theta^{(0)} = \theta^{(t)}$  dan  $\theta = \theta^{(t+1)}$  dengan  $t_{awal} = 0$ , begitu juga untuk  $G$  dan  $H$  sehingga diperoleh iterasi Newton-Raphson sebagai berikut:

$$\theta^{(t+1)} = \theta^{(t)} - (H^{(t)})^{-1} G^{(t)} \tag{2.15}$$

dengan indeks  $t$  menyatakan ukuran iterasi.

Langkah-langkah dari Newton-Raphson sebagai berikut:

- 1) Menentukan estimasi awal dari  $\theta$  yaitu  $\theta^{(0)}$
- 2)  $\hat{\theta}^{(1)} = \hat{\theta}^{(0)} - \frac{G(\hat{\theta}^{(0)})}{H(\hat{\theta}^{(0)})}$ ,  $G(\hat{\theta}^{(0)})$  merupakan turunan pertama dari  $f(\theta)$  pada  $\theta = \hat{\theta}^{(0)}$
- 3)  $\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} - \frac{G(\hat{\theta}^{(t)})}{H(\hat{\theta}^{(t)})}$ , misalkan  $H(\hat{\theta}^{(t)}) = H^{(t)}$  dan  $G(\hat{\theta}^{(t)}) = G^{(t)}$ , maka  $\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} - (H^{(t)})^{-1} G^{(t)}$
- 4) Estimator  $\hat{\theta}^{(t)}$  diiterasi terus sampai diperoleh selisih antara  $\hat{\theta}^{(t+1)}$  dengan  $\hat{\theta}^{(t)}$  nilainya sangat kecil ( $|\hat{\theta}^{(t+1)} - \hat{\theta}^{(t)}| \leq \epsilon$ )

Metode Newton-Raphson dapat diperluas untuk menyelesaikan sistem persamaan dengan lebih dari satu parameter, misalnya  $\theta_1, \theta_2, \dots, \theta_p$  dengan iterasinya sebagai berikut:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - (\mathbf{H}^{(t)})^{-1} \mathbf{G}^{(t)} \quad (2.16)$$

dengan  $\hat{\boldsymbol{\theta}}^{(t+1)} = \begin{pmatrix} \hat{\theta}_1^{(t+1)} \\ \vdots \\ \hat{\theta}_p^{(t+1)} \end{pmatrix}$  dan  $\hat{\boldsymbol{\theta}}^t = \begin{pmatrix} \hat{\theta}_1^{(t)} \\ \vdots \\ \hat{\theta}_p^{(t)} \end{pmatrix}$  sehingga diperoleh

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial^2 \theta_1} & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_1 \theta_2} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_1 \theta_p} \\ \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_1 \theta_2} & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial^2 \theta_2} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_2 \theta_p} \\ \dots & \dots & \ddots & \vdots \\ \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_p \theta_1} & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial \theta_p \theta_2} & \dots & \frac{\partial^2 \ln L(\boldsymbol{\theta})}{\partial^2 \theta_p} \end{pmatrix}$$

dan

$$\mathbf{G} = \begin{pmatrix} \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_1} \\ \vdots \\ \frac{\partial \ln L(\boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix}$$

di mana  $\mathbf{G}$  merupakan turunan pertama dari fungsi *ln-likelihood* dan  $\mathbf{H}$  merupakan turunan kedua dari fungsi *ln-likelihood* yang disebut sebagai matriks Hessian (Nugroho, 2012).

## 2.4 Uji Asumsi Data

### 2.4.1 Kesesuaian Distribusi

Uji kesesuaian distribusi (*goodness of fit*) merupakan pengujian kecocokan atau kesesuaian antara hasil pengamatan (frekuensi pengamatan) tertentu dengan frekuensi yang diperoleh berdasarkan nilai harapannya (frekuensi teoritis). Uji kesesuaian bertujuan untuk mengambil kesimpulan tentang distribusi populasi. Suatu sampel acak dipilih dari populasi bersangkutan, kemudian informasi sampel tersebut digunakan untuk menguji kebenaran distribusi populasi tersebut.

Uji *goodness of fit* untuk mengetahui suatu data berdistribusi Poisson atau tidak dengan menggunakan statistik uji sampel tunggal Kolmogorov-Smirnov. Dalam uji sampel tunggal Kolmogorov-Smirnov pengujian dilakukan pada dua buah fungsi distribusi kumulatif, yaitu distribusi kumulatif yang dihipotesiskan dan distribusi kumulatif yang diamati. Misal diambil sebuah sampel acak dari suatu fungsi distribusi

$F(X)$  yang belum diketahui, akan dipastikan bahwa dapat disimpulkan bahwa  $F(X) = F_0(X)$  untuk semua  $x$ , dengan  $F_0(X)$  adalah fungsi distribusi kumulatif dari suatu sampel acak yang diamati dengan  $N$  pengamatan, dengan  $X$  adalah sembarang nilai yang mungkin, sehingga

$$F_N(X) = \frac{k}{N} \tag{2.8}$$

dengan  $k$  adalah jumlah pengamatan yang sama atau kurang dari  $X$ .

Dalam uji ini diharapkan bahwa untuk setiap  $X$ ,  $F_N(X)$  mendekati  $F_0(X)$ , artinya, di bawah  $H_0$  diharapkan selisih antara  $F(X)$  mendekati  $F_0(X)$  adalah kecil, serta berada dalam batas-batas kesalahan acak.

Uji Kolmogorov-Smirnov memusatkan perhatian pada penyimpangan (deviasi) terbesar. Nilai  $F_0(X) - F_N(X)$  terbesar dinamakan deviasi maksimum dan dinyatakan sebagai berikut:

$$D = maks |F_0(X) - F_N(X)| \tag{2.9}$$

dimana:

$F_0(X)$ : fungsi distribusi kumulatif yang dihipotesiskan

$F_N(X)$ : fungsi distribusi kumulatif yang diamati

#### 2.4.2 Overdispersi

Overdispersi pada Distribusi Poisson dapat dilihat dari nilai dugaan dispersi yaitu nilai *Devians* atau *Pearson chi-square* dibagi dengan derajat bebasnya bernilai lebih besar dari satu. Overdispersi yaitu keadaan nilai variansinya lebih besar dari nilai meannya. Fenomena *overdispersi* dapat ditulis:

$$Var(X) > E[X] \tag{2.10}$$

Dugaan dispersi dapat diukur dengan nilai *Deviance* atau *Pearson chi-square*. Data dikatakan *overdispersi* jika dugaan dispersi lebih besar dari 1 dan *underdispersion* jika dugaan dispersi kurang dari 1 (Khoshgoftaar, et al., 2004).

Terdapat dua cara yang dapat digunakan untuk mendeteksi *overdispersi*, yaitu:

1. *Deviance*

$$\Phi_1 = \frac{D^2}{db}; D^2 = 2 \sum_{i=1}^n \left\{ x_i \ln \left( \frac{x_i}{\lambda} \right) \right\} \tag{2.11}$$

dimana  $db = n - p$  dengan  $p$  merupakan banyak parameter termasuk konstanta,  $n$  merupakan banyaknya pengamatan dan  $D^2$  adalah nilai *Deviance* (Hilbe, 2011).

## 2. *Pearson chi-square*

Untuk menguji asumsi equidispersi pada distribusi Poisson dilakukan dengan melihat nilai statistik *Pearson chi-square* yang dibagi dengan derajat bebasnya.

$$\Phi_2 = \frac{X^2}{db}; X^2 = \sum_{i=1}^n \frac{(x_i - \lambda)^2}{var(x_i)} \quad (2.12)$$

Dimana  $db = n - p$  dengan  $p$  merupakan banyak parameter termasuk konstanta,  $n$  merupakan banyaknya pengamatan dan  $X^2$  adalah nilai *Pearson chi-square*.

Jika nilai  $\Phi_1$  atau  $\Phi_2$  bernilai lebih dari satu maka terjadi overdispersi pada data (Wirajaya Kusuma, 2013).

## 2.5 Pengendalian Kualitas Statistik

Proses pengendalian secara statistik merupakan teknik statistika yang secara luas digunakan untuk memastikan bahwa proses yang sedang berjalan telah memenuhi standar. Pengendalian kualitas statistik adalah salah satu teknik berbeda yang didesain untuk mengevaluasi kualitas ditinjau dari sisi kesesuaian dengan spesifikasinya. Pengendalian kualitas merupakan suatu upaya yang dilaksanakan secara berkesinambungan, sistematis, dan objektif dalam memantau dan menilai barang, jasa, maupun pelayanan dibandingkan dengan standar yang ditetapkan serta menyelesaikan masalah yang ditemukan dengan tujuan untuk menyelidiki dengan cepat terjadinya sebab-sebab terduga atau pergeseran proses sedemikian sehingga dapat dilakukan penyelidikan dan tindakan pembetulan (Hartanti, 2013).

### 2.5.1 Peta Kendali Exponentially Weighted Moving Average

Peta kendali *Exponentially Weighted Moving Average* (EWMA) pertama kali diperkenalkan oleh S.W. Roberts pada tahun 1959. Peta kendali EWMA optimal bagi proses dengan meannya dalam periode  $t$  yang berhubungan dengan mean dalam periode  $(t-1)$ . Peta Kendali EWMA digunakan untuk mendeteksi adanya pergeseran nilai rata-rata yang kecil dalam suatu proses (Montgomery, 2009). Peta kendali EWMA merupakan alternatif peta kendali terhadap peta kendali Shewart dalam mendeteksi pergeseran kecil pada proses mean. Cara kerja peta kendali EWMA yaitu dengan tidak hanya mempertimbangkan sampel terakhir sebagai titik yang diamati, tetapi juga

menggabungkan informasi dalam barisan nilai sampel sebelumnya untuk memperoleh titik observasi yang nantinya akan diamati. Secara khusus EWMA digunakan pada pengamatan secara individu.

Diasumsikan pengamatan dari proses pada variabel  $X \sim N(\mu, \sigma^2)$ . Didefinisikan peta kendali EWMA menurut Montgomery (2009) adalah:

$$Z_t = \theta X_t + (1 - \theta)Z_{t-1} \quad (2.17)$$

dengan  $0 < \theta \leq 1$  adalah konstanta dan nilai awal  $Z_0$  dapat diperoleh berdasarkan nilai rata-rata proses yang diamati sehingga:

$$Z_0 = \mu_0$$

dengan

$Z_t$  = Nilai EWMA pada pengamatan ke- $t$

$\theta$  = parameter pembobot yang bernilai antara 0 dan 1.

$X_t$  = nilai pengamatan ke- $t$ ,  $t = 1, 2, 3, \dots$

$Z_{t-1}$  = rata-rata dari data historis sebelumnya

$Z_0$  = nilai awal

$\mu_0 = \bar{x}$  = rata-0

Sehingga atas kendali atas (UCL) dan batas kendali bawah (LCL) peta kendali EWMA adalah sebagai berikut:

$$\begin{aligned} UCL &= \mu_0 + L\sigma \sqrt{\left(\frac{\theta}{2-\theta}\right) (1 - (1-\theta)^{2t})} \\ CL &= \mu_0 \\ LCL &= \mu_0 - L\sigma \sqrt{\left(\frac{\theta}{2-\theta}\right) (1 - (1-\theta)^{2t})} \end{aligned} \quad (2.18)$$

dengan

$\mu_0$  = rata-rata

$\sigma$  = standar deviasi

$L$  = lebar dari batas kontrolnya

$\lambda$  = parameter bobot

Ketika peta kendali EWMA telah berjalan untuk beberapa periode, persamaan  $(1 - (1 - \lambda)^{2t})$  mendekati satu sehingga batas kendali mendekati nilai keadaan stabil (*steady state*) diberikan oleh persamaan berikut:

$$\begin{aligned}
 UCL &= \mu_0 + L\sigma \sqrt{\left(\frac{\theta}{2-\theta}\right)} \\
 CL &= \mu_0 \\
 LCL &= \mu_0 - L\sigma \sqrt{\left(\frac{\theta}{2-\theta}\right)}
 \end{aligned}
 \tag{2.19}$$

### 2.5.2 Peta Kendali Modifikasi Exponentially Weighted Moving Average

Peta kendali modifikasi EWMA dikembangkan oleh Patel & Divecha pada tahun 2011. Sifat-sifat yang diinginkan dari peta kendali adalah yang mudah untuk menerapkan dan efektif untuk mendeteksi pergeseran dari semua ukuran sesuai spesifikasi. Patel mengkombinasikan peta kendali Shewart dan EWMA yang berguna untuk mendeteksi perubahan kecil dari proses industri yang memiliki data dengan proses autokorelasi berorde satu. Statistik yang diusulkan dengan menambahkan nilai dalam statistik EWMA.

Sifat yang diinginkan dari peta kendali adalah mudah diimplementasikan dan efektif untuk mendeteksi pergeseran semua ukuran sesuai spesifikasi teknis. Peta kendali Modifikasi EWMA mempertimbangkan masa lalu pengamatan yang mirip dengan skema EWMA dan tambahan mempertimbangkan perubahan masa lalu, serta perubahan terbaru dalam suatu proses. Peta kendali modifikasi EWMA menurut Khan, Aslam, dan Jun (2017) adalah sebagai berikut:

$$M_t = \theta X_t + (1 - \theta)M_{t-1} + k(X_t - X_{t-1}) \tag{2.20}$$

dengan

$M_t$  = rata-rata dari data historis

$\theta$  = parameter bobot yang bernilai antara 0 dan 1.

$X_t$  = nilai pengamatan ke - t, t = 1, 2, . .

$X_{t-1}$  = nilai pengamatan sebelumnya

$M_{t-1}$  = rata-rata dari data historis sebelumnya

k = konstanta (parameter tambahan)

Batas kendali atas (UCL) dan batas kendali bawah (LCL) peta kendali modifikasi EWMA adalah sebagai berikut:

$$UCL = \mu_0 + L\sigma \sqrt{\frac{(\theta + 2\theta k + 2k^2)}{(2 - \theta)}}$$

$$CL = \mu_0$$

$$LCL = \mu_0 - L\sigma \sqrt{\frac{(\theta + 2\theta k + 2k^2)}{(2 - \theta)}} \quad (2.21)$$

dengan

$\mu_0$  = rata-rata

$\sigma$  = standar deviasi

$L$  = lebar dari batas kontrolnya

$\theta$  = parameter bobot

$k$  = konstanta (parameter tambahan)

### 2.5.3 Average Run Length

Peta kendali yang lebih cepat mendeteksi sinyal *out of control* disebut lebih sensitif terhadap perubahan proses. Kinerja peta kendali tersebut disebut *Average Run Length* (ARL). ARL adalah banyaknya rata-rata titik atau sampel yang diperlukan sebelum suatu titik atau sampel menyatakan suatu keadaan tidak terkendali. Apabila proses dalam keadaan *in control* maka digunakan  $ARL_0$ . Dengan demikian  $ARL_0$  akan bernilai lebih besar sedangkan  $ARL_1$  akan bernilai lebih kecil apabila dalam keadaan *out of control*. Jika ARL dalam kendali, semakin besar nilai ARL maka semakin baik jenis peta kendali yang bersangkutan. Jika ARL di luar kendali, semakin kecil nilai ARL maka semakin baik jenis peta kendali yang bersangkutan (Montgomery, 2009)

Secara umum persamaan untuk perhitungan ARL adalah sebagai berikut:

$$ARL = \frac{1}{p} \quad (2.22)$$

dengan  $p$  adalah probabilitas suatu titik keluar dari batas-batas peta kendali.

Untuk  $ARL_0$ ,  $p = \alpha$  = probabilitas kesalahan/error tipe I (menyatakan keadaan tidak terkendali padahal keadaan terkendali) atau probabilitas suatu titik rata-rata sampel berada di luar batas kendali pada saat proses terkendali,  $\alpha$  disebut juga sebagai probabilitas *false alarm*. Untuk  $ARL_1$  nilai  $p = 1 - \beta$  = probabilitas kesalahan/error tipe II (menyatakan keadaan terkendali padahal keadaan tidak terkendali) atau probabilitas suatu titik rata-rata sampel berada di dalam batas kendali pada saat proses tidak terkendali. Dimisalkan  $\beta$  adalah peluang bahwa pergeseran proses tidak terdeteksi pada sampel pertama maka peluang bahwa pergeseran proses

terdeteksi pada sampel pertama adalah  $1 - \beta$ . Dengan demikian peluang bahwa pergeseran proses terdeteksi pada sampel ke- $x$  adalah  $\beta^{x-1}(1 - \beta)$ .

Menurut Mashuri dan Rahmawati (2011), jika  $X$  adalah variabel acak yang menyatakan banyaknya sampel sampai ditemukannya *out of control* yang pertama, maka  $P(X = x) = \beta^{x-1}(1 - \beta)$ , sehingga ekspektasi banyaknya sampel yang harus diamati sampai ditemukan *out of control* yang pertama adalah :

$$\begin{aligned} E[X] &= \sum_{x=1}^{\infty} x \beta^{x-1}(1 - \beta) \\ &= (1 - \beta) \sum_{x=1}^{\infty} x \beta^{x-1} \frac{1}{(1 - \beta)^2} = \frac{1}{(1 - \beta)} \end{aligned} \quad (2.23)$$

Oleh karena itu pada kondisi *out of control* adalah:

$$ARL_1 = \frac{1}{(1 - \beta)} \text{ atau } ARL_1 = \frac{1}{P(\text{menolak } H_0 | H_1 \text{ benar})} \quad (2.24)$$

Sedangkan pada kondisi *in control* adalah:

$$ARL_0 = \frac{1}{\alpha} \text{ atau } ARL_0 = \frac{1}{P(\text{menolak } H_0 | H_0 \text{ benar})} \quad (2.25)$$

Bila pengamatan yang diambil dari proses itu saling bebas, maka penentuan ARL mudah ditentukan untuk peta kendali Shewhart karena titik-titik yang diplot pada grafik bersifat saling bebas. Berbeda halnya dengan urutan titik-titik yang diplot pada peta kendali EWMA yang tidak independen, sehingga dibutuhkan pendekatan lain untuk menemukan nilai ARL. Menurut Patel dan Divecha (2011) salah satu metode yang banyak digunakan dalam menentukan nilai ARL pada peta kendali EWMA yaitu metode *markov chain* yang dikembangkan oleh Brook dan Evans (1992). Prosedur dalam metode ini melibatkan pembagian interval di antara nilai UCL dan LCL sebanyak  $N = 2m + 1$  subinterval. Titik tengah  $m_i$  pada subinterval ke- $i$  dapat dituliskan sebagai berikut:

$$m_i = LCL + \frac{(2i - 1)(UCL - LCL)}{2N}$$

dengan subinterval ke- $j$  adalah  $L_j, U_j$  di mana:

$$L_j = LCL + \frac{(j - 1)(UCL - LCL)}{N} \quad (2.26)$$



dan

$$U_j = LCL + \frac{j(UCL - LCL)}{N}$$

Berdasarkan informasi tersebut akan ditentukan matriks peluang transisi dengan menentukan matriks  $Q$ . Elemen matriks  $Q$  untuk baris dan kolom ditandai sesuai dengan banyaknya pengamatan  $N$ . Sehingga matriks  $Q$  berukuran  $N \times N$  berisi:

$$Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} & \dots & Q_{1N} \\ Q_{21} & Q_{22} & Q_{23} & \dots & Q_{2N} \\ Q_{31} & Q_{32} & Q_{33} & \dots & Q_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Q_{N1} & Q_{N2} & Q_{N3} & \dots & Q_{NN} \end{pmatrix}$$

dengan entri  $Q_{ij}$  pada matriks ditentukan dengan cara sebagai berikut:

$$Q_{ij} = P(L_j < Z_t < U_j \mid Z_{t-1} = m_i)$$

$$Q_{ij} = P\left(LCL + \left(\frac{UCL - LCL}{2N\theta}\right)(2j - (1 - \theta)(2i - 1))\right) - P\left(LCL + \left(\frac{UCL - LCL}{2N\theta}\right)(2(j - 1) - (1 - \theta)(2i - 1))\right) \quad (2.27)$$

Sehingga nilai ARL diperoleh dengan rumus sebagai berikut:

$$ARL = (I - Q)^{-1} \mathbf{1} \quad (2.28)$$

dengan

$I$  = Matriks identitas berukuran  $N \times N$

$\mathbf{1}$  = Vektor satuan berukuran  $N \times 1$

dengan hasil dari operasi tersebut akan berupa matriks kolom berukuran  $N \times 1$ , nilai ARL ditentukan berdasarkan median dari elemen yang ada di dalamnya.

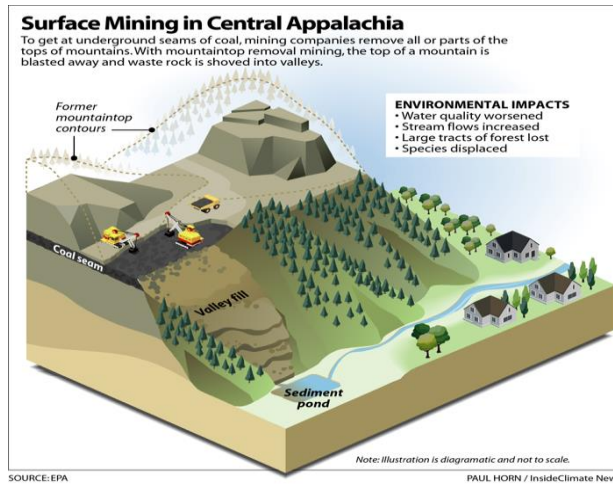
## 2.6 Penambangan/Pemindahan Puncak Gunung & Salamander

Penambangan pemindahan puncak gunung (Mountain Removal Mining/Valley Filling), juga dikenal sebagai pertambangan puncak gunung adalah bentuk dari pertambangan permukaan pada puncak atau puncak punggung gunung yang melibatkan pemindahan sejumlah besar batuan untuk mengakses lapisan batubara yang lebih dalam. Lapisan batubara diekstraksi dari gunung dengan memindahkan

tanah dari permukaan gunung kemudian dipindahkan ke lembah gunung disekitarnya. Penambangan penghilangan puncak gunung (MTR), juga dikenal sebagai penambangan puncak gunung, adalah bentuk dari pertambangan permukaan yang melibatkan perubahan topografi atau penghapusan puncak, bukit, atau punggung bukit untuk mengakses lapisan batu bara yang terkubur.

Ekosistem perairan khususnya terkena dampak karena penambangan pemindahan puncak gunung sering mengakibatkan penguburan secara keseluruhan atau sebagian aliran orde rendah melalui pengisian lembah, proses dimana batuan (yaitu bahan lapisan penutup) dibuang dari lokasi tambang ke lembah yang berdekatan. Secara kolektif, penambangan pemindahan puncak gunung dan penimbunan lembah (MTR/VF) menghasilkan degradasi jangka panjang pada aliran hulu dan daerah sekitarnya (Bernhardt & Palmer 2011). Salamander mewakili vertebrata dominan di sungai tingkat rendah menjadi objek yang terkena dampak dari penambangan dan penimbunan tersebut.

Salamander adalah nama umum bagi sekitar 550 spesies amfibi. Mereka secara umum dicirikan oleh penampilan mirip kadal, dengan tubuh ramping, hidung pendek, dan ekor yang panjang. Semua fosil dan spesiesnya yang telah punah dimasukkan ke dalam ordo *Caudata*, sedangkan spesiesnya yang masih ada digolongkan bersama ke dalam *Urodela*. Sebagian besar salamander memiliki empat jari pada kaki depan mereka dan lima jari pada kaki belakang mereka. Kulit mereka yang lembap menjadikan mereka lebih suka tinggal di habitat di atau dekat air, atau dalam tempat yang terlindung (misalnya tanah lembap), sering kali di lahan basah. Beberapa spesies Salamander sepenuhnya tinggal di air sepanjang hidup mereka, sebagian tinggal di air untuk sementara, dan sebagian lagi tinggal di darat ketika dewasa. Adapun gambaran mengenai proses MRT/VF adalah sebagai berikut:



Sumber: <https://insideclimatenews.org/infographics/infographic-when-mountains-are-leveled-get-their-coal/>

## 2.7 Kerangka Konseptual

