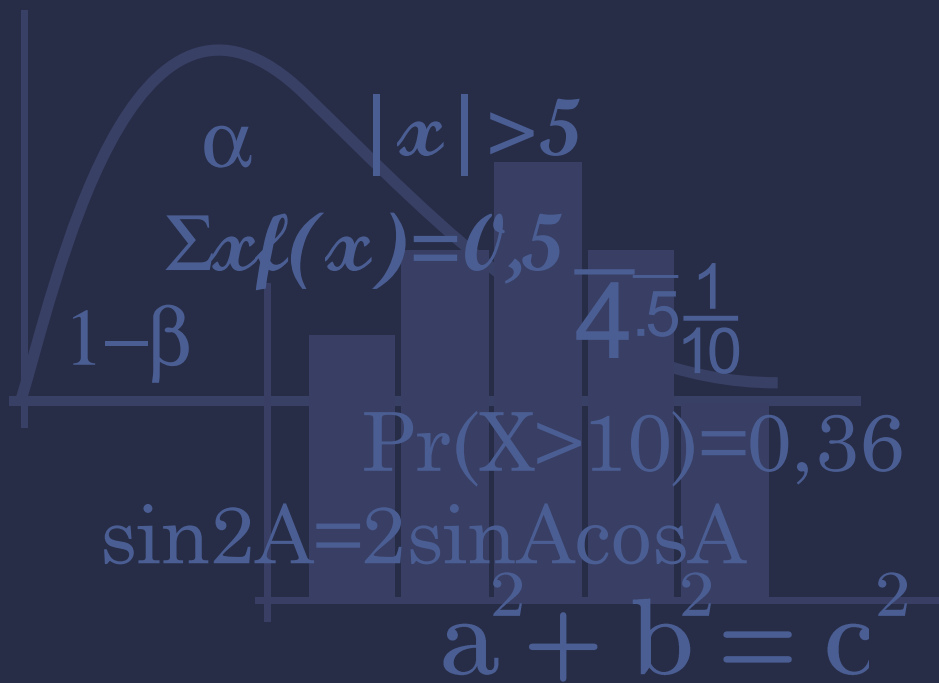


ISBN: 978-602-19356-2-0

prosiding

Seminar Nasional Statistika, Matematika dan Aplikasinya 2014

SNSMA 2014



Dipublikasikan oleh:
Fakultas Matematika & Ilmu Pengetahuan Alan
Universitas Islam Bandung

prosiding

Seminar Nasional Statistika, Matematika dan Aplikasinya 2014

Fakultas Matematika & Ilmu Pengetahuan Alam
Universitas Islam Bandung, Jawa Barat 26 Agustus 2014

**“Statistika dan Matematika
untuk Kemajuan dan Kesejahteraan Umat”**

ISBN: 978 – 602 – 19356 – 2 – 0

Cover Design : Dr. Aceng Komarudin Mutaqin

Tim Prosiding : Dheri Janwar Rusthana, S.Si.
Anjar May Purnama, S.Si.
Octavianty, S.Si.
Fuji Astuti, S.Si.
Maya Setiana, S.Si.
Ihsan Ramadhan, S.Si.

Dipublikasikan oleh : Fakultas Matematika & Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Bandung Jawa Barat

Dewan Editor

Ketua : Dr. Aceng Komarudin Mutaqin

Sekretaris : Suliadi, Ph.D.

Anggota : Prof. Dr. Sutawanir Darwis

Dr. Suwanda

Abdul Kudus, Ph.D.

Dr. Nusar Hajarisman

Dr. Didi Suhaedi

Dr. Yani Ramdhani

Roberta Zulfhi Surya, ST., MT.

Dheri Janwar Rusthana, S.Si.

Kata Pengantar

Puji dan syukur kami panjatkan ke hadirat Alloh SWT, karena hanya dengan izin-Nya maka dapat terselenggara kegiatan SEMINAR NASIONAL STATISTIKA, MATEMATIKA & APLIKASINYA 2014 (SNSMA 2014) oleh Fakultas Matematika & Ilmu Pengetahuan Alam, Universitas Islam Bandung pada hari Selasa, 26 Agustus 2014 di Gedung Pascasarjana UNISBA Jl. Purnawarman No. 59 Bandung. Seminar Nasional Statistika ini bertema “Statistika dan Matematika untuk Kemajuan dan Kesejahteraan Umat.”

Tujuan diadakannya SNSMA 2014 ini adalah dalam rangka diskusi ilmiah, dan tukar menukar informasi di kalangan akademisi, praktisi dan peneliti guna mendorong cepatnya pengembangan ilmu pengetahuan dan teknologi khususnya bidang ilmu statistika dan matematika di Indonesia.

Panitia telah menerima sekitar 75 makalah berasal dari berbagai kalangan, seperti mahasiswa S1, S2, S3, akademisi, praktisi dan peneliti dan berasal dari berbagai daerah di Indonesia. Semua makalah tersebut dipresentasikan pada SNSMA 2014 di Unisba pada Tanggal 26 Agustus 2014, dalam bentuk oral dan dipublikasikan dalam sebuah prosiding. Kami ucapkan terima kasih kepada para peserta pemakalah yang telah berpartisipasi dalam rangka mempercepat pengembangan ilmu pengetahuan dan teknologi.

Semoga Prosiding SNSMA 2014 di Unisba ini bisa bermanfaat dalam penyebarluasan ilmu pengetahuan dan teknologi khususnya bidang bidang ilmu statistika dan matematika di Indonesia. Kepada semua pihak, terutama Tim Prosiding yang telah bekerja keras menyelesaikan prosiding ini, kami ucapkan terima kasih.

Bandung, Agustus 2014

Editor

Daftar Isi

	Halaman
Dewan Editor	i
Kata Pengantar	iii
Daftar Isi	v
Kajian Sosial-Ekonomi Pengembangan Pengolahan Hasil Tangkapan Nelayan Berbasis Masyarakat di Desa Sungai Luar Kec. Batang Tuaka Kab. Indragiri Hilir Riau <i>Ririn Handayani, Hikmatul Hasanah</i>	1-6
Analisis Kekonvergenan Kalman Filter <i>Sutawanir Darwis, Aceng K Mutaqin, Yayat Karyana, Mohammad Sobri</i>	7-17
Analisis Bayesian pada Regresi Binomial dengan Kesalahan Klasifikasi <i>Retno Budiarti</i>	19-26
Analisis Dampak Pemberian Pembiayaan Warung Mikro Bank Syari'ah Mandiri terhadap Usaha Menengah Kecil dan Mikro di Kota Jambi <i>Titin Agustin Nengsih</i>	27-33
Penaksiran Rata-Rata dan Varians dari Distribusi Lognormal pada Data Sampel yang Mengandung Pengamatan Tidak Terdeteksi <i>Dheri Janwar Rusthana, Aceng Komarudin Mutaqin</i>	35-40
Redesign hook-t Muat Sawit untuk Mengurangi Tingkat Keluhan Mosculuskletal Menggunakan Metode Quality Function Deployment (QFD) pada Sentra Pembelian Kelapa Sawit di Pulau Palas Kab. Indragiri Hilir, Riau <i>Rajuli, M.Gasali M, Roberta Zulfhi Surya</i>	41-46
Analisis Antosianin pada Buah Duwet (<i>Syzygium Cumini</i> (L.) Skeels) dengan Metode Ph Differential - Spektrofotometri Sinar Tampak <i>Arlina Prima Putri, Sukanta, Witri Resmisari</i>	47-51
Analisis Penerapan Manajemen Risiko terhadap Profitabilitas Studi pada Bank Umum Syariah di Indonesia Periode Tahun 2011-2013 <i>Okky Paulin</i>	53-59

Analisis Customer Gap dengan Metode Servqual (Service Quality) di Restoran Dapur Iga Bandung <i>Siti Fadilah Ristekawati</i>	61-70
Identifikasi Faktor-Faktor yang Mempengaruhi Angka Kematian Bayi di Provinsi Jawa Barat melalui Model Berbasis Regional <i>Nusar Hajarisman</i>	71-84
Penduga Area Kecil pada Model Level Area <i>Erwin Tanur</i>	85-91
Model Zero-Inflated dan Hurdle pada Data Hitung dengan Banyak Respon Nol <i>Erwin Tanur</i>	93-98
Studi Keamanan Bahan Kimia Obat dan Pangan Berbasis Software <i>Diar Herawati</i>	99-107
Pengembangan Indikator Strip Berbasis Komposit Poli (Metilmetakrilat) - Polisulfonat untuk Identifikasi Formalin pada Sampel Makanan <i>Arlina Prima Putri, Anggi Arumsari, Tutuh Maftuhah</i>	109-112
Analisis Asam Retinoat pada Krim Pemutih Wajah Menggunakan Kromatografi Lapis Tipis dan Kromatografi Cair Kinerja Tinggi <i>Arlina Prima Putri, Sukanta, Adinda A. Nastiti</i>	113-118
Aktivitas Antioksidan Fraksi Ekstrak Etanol Daun Jambu Air [Eugenia Aqueum (Burn.F) Alston] Secara <i>In Vitro</i> dengan Metode Carotene Bleaching <i>Suwendar dan Siti Hazar</i>	119-124
Perbandingan Model Statistika bagi Penentuan Batas Kritis Hara Fosfor pada Kedelai <i>Mohammad Masjkur dan Wiwik Hartatik</i>	125-130
Ruang Modular <i>Mariatul Kiftiah, Yundari</i>	131-139
Manajemen Shift Kerja Berdasarkan Biaya Tenaga Kerja pada PT. XYZ Palm Mill <i>Roberta Zulfhi Surya</i>	141-145
Hubungan antara Penilaian Pengguna Jalan terhadap Kondisi <i>Traffick Light</i> dengan Kenyamanan Berkendara pada Persimpangan Lampu Merah Batang Tuaka Tembilahan <i>M. Gasali, M, Akbar Alfa</i>	147-152

Pemodelan Regresi Logistik Multinomial untuk Data Asuransi Jiwa Dwiguna pada PT. XXX	
<i>Ihsan Ramadhan, Aceng Komarudin Mutaqin, Lisnur Wachidah</i>	153-161
Pengaruh Tingkat Kesehatan Bank terhadap Harga Saham Tahun 2009-2013	
<i>Tri Indriyani</i>	163-170
Kelayakan Teknis Pembangunan Pelabuhan Pendaratan Ikan di Kuala Enok Kabupaten Indragiri Hilir Riau	
<i>Akbar Alfa, Masykur HZ.</i>	171-176
Kelayakan Sosial – Ekonomi Pembangunan Pelabuhan Pendaratan Ikan di Kuala Enok Kabupaten Indragiri Hilir Riau	
<i>Ririn Handayani, Hikmatul Hasanah</i>	177-183
Penggunaan Pemrograman Dinamik pada Pengalokasian Pertambahan Server Pelayanan Kesehatan	
<i>Elis Ratna Wulan dan Isna Lathifah</i>	185-191
Pemetaan Karakteristik Kabupaten/Kota di Provinsi Jawa Barat Berdasarkan Indikator Pembangunan Sosial dan Ekonomi menggunakan Metode Biplot	
<i>Andie Rossandi, Suliadi, Siti Sunendiari, Kurdi</i>	193-199
Pemodelan <i>Vector Autoregression</i> Bivariat untuk Meramalkan Inflasi dan Suku Bunga Sertifikat Bank Indonesia Berdasarkan Data Periode Juni 2006 – Mei 2014	
<i>Maya Setiana, Sutawanir Darwis, Siti Sunendiari</i>	201-211
Pendekatan Kontekstual untuk Meningkatkan Kemampuan Pemecahan Masalah Matematis dan Kemandirian Belajar Siswa SMP	
<i>Ratna Sariningsih</i>	213-220
Penerapan Pendekatan Generatif terhadap Kemampuan Pemecahan Masalah Matematik Siswa SMP	
<i>Nelly Fitriani</i>	221-225
Metode <i>Collaborative Learning</i> untuk Meningkatkan Kemampuan Komunikasi Matematis Siswa SMP	
<i>Anik Yuliani</i>	227-231
Perbandingan Mengajar Dua Guru yang Berbeda Wilayah di SMA XYZ	
<i>Siti Sunendiari, Endah Kusumastuti</i>	233-238
Perbandingan Peringkat Efisiensi Kinerja Reksa Dana Syariah dan Konvensional menggunakan Data Envelopment Analysis (DEA)	
<i>Laela Tri Nur Ilaina</i>	239-248

Model Cure dengan Distribusi Peluang Log-logistik untuk Pemodelan Debitur Macet	<i>Fuji Astuti, Abdul Kudus, Lisnur Wachidah</i>	249-258
Model Komitmen Organisasi, Motivasi dan Pembinaan Sumber Daya Manusia (SDM) terhadap Kinerja Pegawai (Studi pada Pemerintah Kabupaten Oku Timur)	<i>Anuar Sanusi dan Yulmaini</i>	259-266
Optimasi Biaya Total Persediaan dengan Permintaan Bersifat Linier	<i>Muhammad Ghani Fathurrahman, M. Yusuf Fajar, Yani Ramdani</i>	267-271
Analisis Perilaku Berbahaya yang Dominan pada Pengendara Sepeda Motor di Kota Tembilahan	<i>Siti Nurkamila Insani, Akbar Alfa, M. Gasali, M.</i>	273-277
Sistem Pakar Diagnosa Kerusakan pada Televisi Berwarna dengan Metode Forward Chaining (Studi Kasus: Jurusan Audio Video SMK 2 Tembilahan)	<i>Agustriyan dan Dwi Yuli Prasetyo</i>	279-295
Pengukuran Kewajaran Harga Saham yang Mendekati Batas Bawah Aturan Autorejection di Bursa	<i>Muhammad Rifqi Syauqi</i>	297-304
Aktivitas Hepatoprotektif Ekstrak Remis (<i>Corbicula Javanica Mousson</i>) terhadap Tikus Putih Jantan Galur Wistar	<i>Faza Shalihah Novani, Sri Peni Fitriarningsih, Siti Hazar</i>	305-310
Efektivitas Penerapan Regresi Linier Berganda Dua Prediktor pada Kajian Data Survei Berbasis Skala Sikap Likert	<i>Soekardi Hadi Prabowo</i>	311-317
Model <i>Geographically Weighted Regression</i> pada Analisis Tingkat Kemiskinan Kabupaten/Kota di Provinsi Jawa Barat	<i>Octavianty, Siti Sunendiari, Teti Sofia Yanti</i>	319-328
Kesalahan Tipe I Empirik Score Test untuk Poisson Mixed Model Melawan Zero Inflated Poisson Mixed Model	<i>Nurvita Trianasari</i>	329-335
Kuasa Uji Likelihood Ratio Test untuk Model Regresi Poisson Melawan Model Regresi Zero-Inflated Poisson	<i>Nurvita Trianasari</i>	337-342

Value-at-Risk Contribution Di Bawah Model Aset Liabilitas Menggunakan Pendekatan EWMA <i>Sukono, Sudradjat Supian, Dwi Susanti</i>	343-352
Sengkala-Sengkala dari Asia Tenggara yang Melawan Prinsip <i>Angkanam Vamato Gatih</i> <i>Agung Prabowo</i>	353-362
Pengklastran Data dengan Menggunakan <i>Divisive Analysis Method (DIANA)</i> <i>Chandra Gunawan, Dewi Rachmatin, dan Maman Suherman</i>	363-374
Kinerja Model Generalized Poisson Regression dan Negative Binomial Regression dalam Mengatasi Masalah Over/under dispersion pada Model Regresi Data Count <i>Cucu Sumarni</i>	375-383
Analisis Kestabilan Model Penyebaran dan Pengendalian Penyakit Tuberculosis <i>Embay Rohaeti, Sri Wardatun dan Ani Andriyati</i>	385-391
Optimasi Biaya Pendistribusian Minyak Tanah dengan Metode Trasportasi (Studi Kasus : PT. Pertamina Kab. Sanggau KALBAR) <i>Bayu Prihandono, Fajria Aryanti, Beni Irawan</i>	393-398
Metode Proyeksi Biproporsional untuk Melihat Perubahan Struktur Ekonomi dengan Memanfaatkan Tabel Input-Output <i>Anugerah Karta Monika</i>	399-407
Pengembangan Model Pembelajaran Mata Kuliah Pengantar Aljabar Abstrak dengan <i>Clustering</i> Mahasiswa Berdasarkan Nilai Prestasi Belajar Modul Prasyarat Menggunakan Algoritma <i>Fuzzy C- Means</i> <i>Nilamsari Kusumastuti</i>	409-417
Pengembangan Indikator Strip Formalin Berbasis Poli METILMETAKRILAT)-Polisulfon-Silika Gel untuk Pemeriksaan Formalin pada Makanan <i>Arlina Prima Putri, Sukanta, Averoer Prabowo</i>	419-423
Analisis Pendapatan dan Efisiensi Usaha Petani Pemilik Penggilingan Padi Kecil (Studi Kasus di Desa Boros Kecamatan Tanjungkerta Kabupaten Sumedang) <i>Diyani Fauziyah</i>	425-434
Analisis Trend untuk Meramalkan Nilai PDRB Kota Bandung <i>Teti Sofia Yanti dan Onoy Rohaeni</i>	435-441

Membanding Penggunaan LISREL dan SPSS dalam Analisis Jalur <i>Suparman Ibrahim Abdullahi, Maria Cleopatra, Sara Sahrazad</i>	443-450
Kalender Masehi Kembar <i>Riyanto</i>	451-454
Meningkatkan Kemampuan Berpikir Kritis Matematis Siswa SMP Melalui Model Pembelajaran <i>Guided Teaching</i> <i>In Hi Abdullah dan Isman Muhammad Nur</i>	455-464
Aplikasi Var dalam Analisis Risiko pada Porto Folio Single Index Model (Studi Kasus: Data Indeks Harga Saham Jakarta Islamic Index) <i>Edi Saputra, Neva Satyahadewi, Evy Sulistianingsih</i>	465-474
Kemampuan Guru Matematika dalam Mengintegrasikan Aplikasi Teknologi Informasi dan Komunikasi pada Pendekatan Saintifik Guna Mendukung Implementasi Kurikulum 2013 <i>Euis Eti Rohaeti</i>	475-480
Pendekatan Analisis Kelompok untuk Mengelompokkan Desa di Kabupaten Sleman Berdasarkan Tingkat Kerentanan terhadap Bencana <i>Siti Arni Wulandya dan Akhmad Fauzy</i>	481-486
Aplikasi Metode Control Chart dan Regresi Linier Berganda pada Burner Temperature Sistem Thermal Oxidizer di PT. MEDCO E&P Field Singa Lematang, Sumatera Selatan <i>Zamzam Muntaz dan Akhmad Fauzy</i>	487-494
Kajian Simulasi Tingkat Kepercayaan Bagi Parameter, Fungsi Tahan Hidup dan Kuantil Waktu Hidup Dari Data Berdistribusi Eksponensial Dua Parameter Tersensor Tipe-II Double <i>Akhmad Fauzy</i>	495-502
Deskripsi Luas Panen, Produksi dan Produktivitas Padi Sawah dan Padi Ladang di Kabupaten Lombok Timur <i>Lalu Asri Adhitya Nugraha dan Akhmad Fauzy</i>	503-510
Analisis Spasial Kerentanan Sosial Kabupaten Sleman Menggunakan Data Sensus Penduduk 2010 <i>Riswan Dwiramadhan dan Akhmad Fauzy</i>	511-520

Penaksiran Rata-Rata dan Varians dari Distribusi Lognormal pada Data Sampel yang Mengandung Pengamatan Tidak Terdeteksi

Dheri Janwar Rusthana, Aceng Komarudin Mutaqin

Program Studi Statistika Universitas Islam Bandung

Email : dherijr@gmail.com

ABSTRAK

Pengamatan tidak terdeteksi sering kali dihadapi dalam menganalisis data lingkungan yang nilainya terlalu kecil untuk dapat diukur, sehingga dicatat sebagai suatu pengamatan tidak terdeteksi. Zhong (2005) membahas penaksiran rata-rata dan varians populasi distribusi lognormal untuk data sampel yang mengandung pengamatan tidak terdeteksi menggunakan metode PKM (Penaksiran Kemungkinan Maksimum) yang diintegrasikan dengan algoritme EM (Ekspektasi-Maksimisasi). Dalam makalah ini diusulkan metode lain untuk kasus di atas dengan menggunakan penaksir PKM disesuaikan yang bersifat tak bias dan bervarians minimum yang diintegrasikan dengan algoritme EM. Data sekunder digunakan untuk mengevaluasi kinerja dari metode usulan di atas dengan mengkondisikan ke dalam beberapa kasus data. Hasilnya menunjukkan bahwa metode yang diusulkan dalam makalah ini lebih baik dibandingkan dengan metode Zhong.

Kata Kunci : Penaksir Kemungkinan Maksimum, Algoritme EM, Distribusi Lognormal, Absolute Relative Bias, Mean Squared Error.

1. PENDAHULUAN

Adanya pengamatan yang tidak terdeteksi sering kali dihadapi dalam menganalisis data lingkungan, misalnya konsentrasi suatu zat yang nilainya terlalu kecil untuk dapat diukur, sehingga dicatat sebagai suatu pengamatan tidak terdeteksi. Masalah penaksiran parameter berdasarkan data yang mengandung pengamatan tidak terdeteksi sangat menarik untuk diteliti. Pengamatan yang tidak terdeteksi tidak dapat dengan mudah dibuang tanpa mengakibatkan adanya masalah bias dalam penaksiran parameter.

Pendekatan yang biasa digunakan untuk menaksir parameter populasi berdasarkan data sampel yang mengandung pengamatan tidak terdeteksi adalah metode substitusi, metode parametrik dan nonparametrik. Metode substitusi mengganti pengamatan tidak terdeteksi dengan suatu angka yang nilainya tergantung pada kemampuan alat ukur yang disebut dengan Batas Deteksi (BD). Umumnya praktisi mengganti pengamatan tidak terdeteksi dengan $BD/2$ atau nilai nol (0) kemudian menggunakan metode statistika untuk data lengkap. Tidak ada alasan rasional mengganti pengamatan tidak terdeteksi dengan cara di atas.

Pendekatan parametrik mengasumsikan data berasal dari suatu populasi yang berdistribusi tertentu. Jika ukuran sampelnya cukup dan distribusi yang dipilih hampir sama dengan distribusi data aktual, maka penaksiran kemungkinan maksimum bersifat efisien. Tetapi mungkin saja syarat tersebut tidak terpenuhi, sehingga penaksirannya tidak lagi bersifat efisien. Gleit (1985) menunjukkan bahwa pendekatan parametrik mempunyai kinerja yang buruk untuk data sampel berukuran 25 sampai 50.

Pendekatan nonparametrik tidak mengasumsikan apapun mengenai bentuk distribusi datanya dan kinerja penaksir nonparametrik ini cukup baik untuk ukuran sampel kecil ($n < 50$) dan persentase pengamatan tidak terdeteksinya dalam tingkat yang sedang (Gilbert, 1987).

Zhong (2005) membahas penaksiran rata-rata dan varians populasi untuk data sampel yang mengandung pengamatan tidak terdeteksi. Data sampelnya diasumsikan berasal dari suatu populasi yang berdistribusi lognormal. Penaksiran parameter dan pengamatan tidak terdeteksi dicari dengan menggunakan metode PKM (Penaksiran Kemungkinan Maksimum) yang diintegrasikan dengan algoritme EM (Ekspektasi-

Maksimisasi). Taksiran rata-rata dan varians populasinya diperoleh dari penggunaan metode PKM dengan mengganti parameter-parameter yang ada pada rata-rata dan varians dari distribusi lognormal oleh taksiran parameter-parameternya.

Shen (1998) menunjukkan bahwa taksiran PKM untuk rata-rata dan varians dari distribusi lognormal adalah bersifat bias. Finney (1941) menurunkan penaksir PKM disesuaikan yang bersifat tak bias dan bervarians minimum untuk rata-rata dan varians dari distribusi lognormal. Hasil yang sama juga ditunjukkan oleh Shen (1998) namun dengan cara yang berbeda. Tujuan dari makalah ini adalah untuk mendapatkan metode penaksiran rata-rata dan varians dari populasi yang berdistribusi lognormal dimana data sampelnya mengandung pengamatan tidak terdeteksi yang mengintegrasikan algoritme EM dengan PKM yang disesuaikan hasil dari Finney (1941). Tujuan selanjutnya adalah untuk mengetahui perbedaan kinerja dari metode usulan dengan metode yang digunakan oleh Zhong (2005) berdasarkan kriteria bias relatif absolut (*Absolute Relative Bias – ARB*) dan rata-rata kuadrat kekeliruan (*Mean Squared Error – MSE*) untuk kasus-kasus data yang ditentukan.

2. TINJAUAN PUSTAKA

2.1. Distribusi Lognormal

Distribusi dari data lingkungan sering kali miring ke kanan dan biasanya dimodelkan oleh distribusi lognormal. Distribusi lognormal mempunyai fungsi densitas peluang sebagai berikut:

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

Untuk $X > 0$, maka $Y = \ln X$ akan berdistribusi normal dengan rata-rata μ dan varians σ^2 . Fungsi distribusi kumulatif untuk distribusi lognormal adalah:

$$F(x) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right) \quad (1)$$

dimana $\Phi(\cdot)$ adalah notasi untuk fungsi distribusi kumulatif normal baku. Rata-rata dan varians dari distribusi lognormal masing-masing adalah:

$$\theta = e^{\mu + \frac{\sigma^2}{2}} = \exp\left(\mu + \frac{1}{2}\sigma^2\right) \quad (2)$$

$$\eta^2 = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1] \quad (3)$$

Penaksir kemungkinan maksimum untuk θ dan η^2 dari distribusi lognormal adalah:

$$\hat{\theta} = \exp\left(\hat{\mu} + \frac{1}{2}\hat{\sigma}^2\right) \quad (4)$$

$$\hat{\eta}^2 = \exp(2\hat{\mu} + \hat{\sigma}^2)[\exp(\hat{\sigma}^2) - 1] \quad (5)$$

dimana

$$\hat{\mu} = \frac{\sum_{i=1}^n \ln X_i}{n}, \text{ dan} \quad (6)$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (\ln X_i - \hat{\mu})^2}{n} \quad (7)$$

2.2. PKM TERINTEGRASI ALGORITME EM

Maximum likelihood estimation (MLE) atau metode pendugaan kemungkinan maksimum (PKM) adalah metode parametrik yang mengasumsikan bahwa data pengamatan berasal dari suatu populasi yang berdistribusi tertentu.

Misalkan x_i menyatakan pengamatan terdeteksi ke- i , dengan $i = 1, 2, \dots, n$; p menyatakan banyaknya jenis alat ukur dengan BD yang berbeda-beda, dan t_j menyatakan banyaknya pengamatan tidak terdeteksi untuk BD_j , dengan $j = 1, 2, \dots, p$, maka rumus umum fungsi kemungkinan untuk data tidak terdeteksi dengan berbagai BD adalah:

$$L(\theta_1, \theta_2, \dots, \theta_k) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k) \left\{ \prod_{j=1}^p \left(\prod_{m=1}^{t_j} F(BD_j | \theta_1, \theta_2, \dots, \theta_k) \right) \right\}$$

dengan $\theta_1, \theta_2, \dots, \theta_k$ adalah parameter-parameter distribusi, f adalah fungsi densitas peluang dan F adalah fungsi distribusi kumulatif. Untuk memudahkan komputasi, sering kali digunakan fungsi log-kemungkinan (*log likelihood*). Untuk distribusi lognormal, fungsi log-kemungkinan-nya adalah:

$$l(\mu, \sigma) = \sum_{j=1}^p t_j \ln \left[\Phi \left(\frac{\ln(BD_j) - \mu}{\sigma} \right) \right] - \frac{n}{2} \ln[2\pi\sigma^2] - \frac{1}{2} \sum_{i=1}^n \frac{(\ln(x_i) - \mu)^2}{\sigma^2},$$

nilai μ dan σ^2 yang memaksimumkan fungsi di atas merupakan penaksir PKM bagi μ dan σ^2 .

Algoritme EM adalah metode iteratif yang digunakan untuk mendapatkan nilai taksiran kemungkinan maksimum untuk data tidak lengkap. Metode ini terdiri dari atas cara untuk menaksir pengamatan tidak lengkap berdasarkan nilai taksiran parameter sementara (langkah Ekspektasi, E) dan menaksir parameter berdasarkan hasil pengamatan aktual dan pengamatan hasil taksiran (langkah Maksimisasi, M).

Zhong (2005) menggunakan metode PKM yang diintegrasikan dengan algoritme EM untuk menaksir rata-rata dan varians populasi untuk data sampel yang mengandung pengamatan tidak terdeteksi, dimana data sampelnya diasumsikan berasal dari suatu populasi yang berdistribusi lognormal. Dalam langkah E dari metode Zhong ini, data tidak terdeteksi diganti oleh nilai ekspektasi bersyaratnya yang sesuai dengan batas deteksi BD_j , untuk $j = 1, 2, \dots, p$. Untuk kasus distribusi lognormal, langkah ini sama halnya dengan mengganti logaritma dari data tidak terdeteksi dengan:

$$E(\ln(X_i) | \ln(X_i) < \ln(BD_j)) = \mu - \sigma \left[\frac{\phi \left(-\frac{\ln(BD_j) - \mu}{\sigma} \right)}{1 - \Phi \left(-\frac{\ln(BD_j) - \mu}{\sigma} \right)} \right] \quad (8)$$

dengan ϕ dan Φ masing-masing merupakan fungsi densitas dan fungsi distribusi kumulatif normal baku. Nilai awal bagi μ dan σ^2 adalah penaksir PKM untuk μ dan σ^2 dari pengamatan-pengamatan langkah E yang pertama. Dalam langkah M dari metode Zhong, penaksir $\hat{\mu}$ dan $\hat{\sigma}^2$ diperoleh dengan memaksimumkan fungsi log-kemungkinan dari data pengamatan terdeteksi yang digabungkan dengan nilai dugaan pengamatan tidak terdeteksi yang diperoleh pada langkah E. Langkah E dan M tersebut terus diulang hingga konvergen. Taksiran rata-rata dan varians populasi untuk data sampel yang mengandung pengamatan tidak terdeteksi diperoleh dengan mensubstitusikan nilai $\hat{\mu}$ dan $\hat{\sigma}^2$ pada Persamaan (4) dan (5).

3. PKM DISESUAIKAN YANG TERINTEGRASI ALGORITME EM

Shen (1998) menunjukkan bahwa penaksir dengan menggunakan PKM untuk θ dan η^2 dari distribusi lognormal adalah bias karena

$$E(\hat{\theta}) = \theta \left(1 - \frac{\sigma^2}{n} \right)^{-\frac{(n-1)}{2}} \exp \left\{ -\frac{n-1}{n} \left(\frac{\sigma^2}{2} \right) \right\}, \text{ dan}$$

$$E(\hat{\eta}^2) = \eta^2 \left\{ \left(1 - \frac{4\sigma^2}{n} \right)^{-\frac{(n-1)}{2}} - \left(1 - \frac{2\sigma^2}{n} \right)^{-\frac{(n-1)}{2}} \right\} \exp \left\{ \left(\frac{2}{n} - 1 \right) \sigma^2 \right\} \{ \exp(\sigma^2) - 1 \}^{-1}.$$

Finney (1941) mendefinisikan suatu deret tak hingga

$$g(t) = 1 + t + \frac{n-1}{n+1} \left(\frac{t^2}{2!} \right) + \frac{(n-1)^2}{(n+1)(n+3)} \left(\frac{t^3}{3!} \right) + \dots \quad (9)$$

dan mendapatkan penaksir kemungkinan maksimum yang disesuaikan untuk θ dan η^2 dari distribusi lognormal yaitu:

$$\tilde{\theta} = \exp(\hat{\mu}) g\left(\frac{1}{2}\hat{\sigma}^2\right), \quad (10)$$

dan

$$\tilde{\eta}^2 = \exp(2\hat{\mu}) \left\{ g(2\hat{\sigma}^2) - g\left(\frac{n-2}{n-1}\hat{\sigma}^2\right) \right\} \quad (11)$$

yang masing-masing bersifat tak-bias.

Dalam makalah ini diusulkan metode penaksiran rata-rata dan varians untuk data yang mengandung pengamatan tidak terdeteksi yang berasal dari distribusi lognormal dengan memanfaatkan hasil dari Finney (1941) di atas. Taksiran rata-rata dan varians populasi untuk data sampel yang mengandung pengamatan tidak terdeteksi diperoleh dengan mensubstitusikan nilai $\hat{\mu}$ dan $\hat{\sigma}^2$ pada Persamaan (10) dan (11), dimana nilai $\hat{\mu}$ dan $\hat{\sigma}^2$ diperoleh secara iteratif sebagaimana pada bagian sebelumnya.

4. DATA

Metode usulan dalam makalah ini dan juga metode yang digunakan oleh Zhong akan dicoba penerapannya pada data konsentrasi Atrazin (Atz) yang ada pada Junk (1980) dalam Hesel (2012) yang tertera dalam Tabel 1. Atrazine merupakan obat pembunuh rumput yang paling umum digunakan. Atrazine juga dapat mengubah katak jantan jadi betina (perubahan genetika) ketika tingkat konsentrasinya (kandungan) di dalam air terhitung besar.

Tabel 1 menyajikan data konsentrasi Atrazin dalam air tanah dalam satuan mikrogram/liter (mcg/L) di 19 sumur pada bulan September. Untuk melihat kinerja dari metode yang diusulkan, pada data tersebut akan diterapkan beberapa kasus batas deteksi (BD). Secara umum BD yang digunakan adalah 0,10, 0,07, 0,06, 0,04, dan 0,03.

Tabel 1 Konsentrasi Atrazin dalam air tanah (mcg/L) di 19 sumur.

Sumur	Konsentrasi
(1)	(2)
1	2,26
2	0,63
3	0,59
4	0,05
5	0,84
6	0,58
7	0,02
8	0,09
9	0,31
10	0,02
11	0,50
12	0,02
13	0,09
14	0,06
15	0,03
16	0,10
17	0,25
18	0,03
19	88,36

Kolom (1) pada Tabel 1 merupakan nomor sumur yang dijadikan sebagai -objek penelitian, sedangkan kolom (2) merupakan data Konsentrasi pada masing-masing sumur.

Tabel 2 berisikan kasus-kasus data untuk membandingkan metode Zhong dengan metode yang diusulkan dalam makalah ini untuk menaksir rata-rata dan varians dari distribusi lognormal yang data sampelnya mengandung pengamatan tidak terdeteksi.

Tabel 2 Kasus data yang digunakan dalam penelitian

Smr.	Data Asli	Kasus Data											
		1	2	3	4	5	6	7	8	9	10	11	12
1	0,02	< 0,10	< 0,07	< 0,06	< 0,04	< 0,03	< 0,03	< 0,03	< 0,03	< 0,03	0,02	0,02	0,02
2	0,02	< 0,10	< 0,07	< 0,06	< 0,04	< 0,03	< 0,03	< 0,03	< 0,03	< 0,03	0,02	0,02	0,02
3	0,02	< 0,10	< 0,07	< 0,06	< 0,04	< 0,03	0,02	0,02	0,02	0,02	0,02	0,02	0,02
4	0,03	< 0,10	< 0,07	< 0,06	< 0,04	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03
5	0,03	< 0,10	< 0,07	< 0,06	< 0,04	0,03	< 0,04	< 0,04	< 0,04	0,03	< 0,04	< 0,04	0,03
6	0,05	< 0,10	< 0,07	< 0,06	0,05	0,05	< 0,06	< 0,06	0,05	< 0,06	< 0,06	< 0,06	< 0,06
7	0,06	< 0,10	< 0,07	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06
8	0,09	< 0,10	0,09	0,09	0,09	0,09	< 0,10	0,09	0,09	< 0,10	< 0,10	0,09	< 0,10
9	0,09	< 0,10	0,09	0,09	0,09	0,09	< 0,10	0,09	0,09	< 0,10	< 0,10	0,09	< 0,10
10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10	0,10
11	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25	0,25
12	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31
13	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50
14	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58	0,58
15	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59	0,59
16	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63	0,63
17	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84
18	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66	2,66
19	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36	88,36

5. HASIL DAN PEMBAHASAN

Data konsentrasi Atrazin dalam air tanah berasal dari populasi yang berdistribusi lognormal, hal ini didasarkan pada uji Anderson-Darling yang nilai statistik ujinya $A_{19}^2 = 0,6213$ sedangkan nilai kritisnya adalah 3,857 (untuk taraf nyata $\alpha = 5\%$).

Tabel 3 menyajikan perhitungan nilai ARB dan MSE untuk penaksiran parameter rata-rata dan varians dari distribusi lognormal menggunakan metode Zhong dan metode usulan baik untuk kasus 1 dan 2. Kasus 1 adalah ketika nilai parameter sesungguhnya dari rata-rata dan varians distribusi lognormal diperoleh melalui metode PKM dari data asli pada Tabel 1. Sedangkan kasus 2 adalah ketika nilai parameter sesungguhnya dari rata-rata dan varians distribusi lognormal diperoleh melalui metode PKM yang disesuaikan dari data asli pada Tabel 1. Terlihat bahwa nilai-nilai ARB dan MSE untuk metode usulan lebih kecil dibandingkan dengan nilai-nilai ARB dan MSE untuk metode Zhong baik untuk penaksiran rata-rata maupun varians. Hal ini menunjukkan bahwa untuk kasus-kasus data yang ada pada Tabel 2, metode usulan memiliki kinerja yang lebih baik dibandingkan dengan metode Zhong dalam menaksir parameter rata-rata dan varians dari distribusi lognormal.

Tabel 3 Nilai ARB dan MSE

Kasus	Parameter	ARB		MSE	
		Zhong	Usulan	Zhong	Usulan
1	Rata-rata	0,3154	0,0480	0,2514	0,0080
	Varians	4,8147	0,6252	573.330,7006	6.627,0703
2	Rata-rata	0,2344	0,0603	0,1674	0,0133
	Varians	9,1301	0,3556	629.741,5897	800,1262

6. KESIMPULAN DAN SARAN

Dalam makalah ini telah dirumuskan metode penaksiran rata-rata dan varians dari populasi yang berdistribusi lognormal dimana data sampelnya mengandung pengamatan tidak terdeteksi yang mengintegrasikan algoritme EM dengan PKM yang disesuaikan hasil dari Finney (1941). Hasil perbandingan antara metode usulan di atas dengan metode yang telah ada yaitu metode PKM yang diintegrasikan dengan algoritme EM milik Zhong (2005) untuk kasus-kasus data yang dicobakan dalam makalah ini menunjukkan bahwa metode usulan memiliki kinerja yang lebih baik dibandingkan dengan metode Zhong. Hal ini terlihat dari nilai-nilai ARB dan MSE dari metode usulan lebih kecil dibandingkan nilai-nilai ARB dan MSE dari metode Zhong.

Beberapa saran yang dapat dikemukakan dalam makalah ini yang dianggap penting untuk diperhatikan yaitu: (1) bagi para pengguna statistika disarankan untuk menggunakan metode PKM disesuaikan yang diintegrasikan algoritme EM untuk menaksir parameter rata-rata dan varians dari distribusi lognormal dimana data sampelnya mengandung pengamatan tidak terdeteksi dikarenakan penaksiran parameter rata-rata dan varians dari metode yang digunakan oleh Zhong (2005) bersifat bias; (2) Bagi peneliti lain disarankan untuk membuat program yang lebih efisien dalam membangun program untuk menghitung nilai deret yang ada dalam metode PKM disesuaikan karena dalam program yang digunakan untuk skripsi ini hanya terbatas pada 20 suku deret; (3) Bagi peneliti lain disarankan untuk meneliti lebih lanjut dalam membandingkan metode Zhong dan metode usulan pada skripsi dengan menggunakan simulasi.

DAFTAR PUSTAKA

- Finney, D. J., (1941). On The Distribution of A Variate whose Logarithm is Normally Distributed. *Royal Statistical Society Journal*, 7, B, 155-161.
- Gilbert, R. O., (1987). *Statistical Method for Environmental Pollution Monitoring*, New York. Wiley.
- Gleit, A., (1985). Estimation for Small Normal Data Sets with Detection Limits. *Environmental Science Ana Technology*, 19, 1201-1206.
- Hessel, R. (2012). *Statistics for Censored Environmental Data Using Minitab And R*. New York. Wiley.
- Klugman, A., Panjer, H., Willmot, E., (2008). *Loss Models krom Data do Decisions*. Edisi Ketiga. New York. Wiley.
- Law, A. M., Kelton, W. D., (1991). *Simulation Modeling and Analysis*. Edisi Kedua, McGraw-Hill International Edition, 337.
- Shen, W. H., (1998). Estimation of Parameters of A Lognormal Distribution, *Taiwanese Journal of Mathematics*, 243-250.
- Zhong, W. (2005). *Statistical Approaches to Analyze Censored Data with Multiple Detection Limits*. Disertasi Program Doctor of Philosophy Universitas Cincinnati.