

Volume 1, Nomor 1, Tahun 2013

ISSN 2338-7122

Lhokseumawe,
28-29 AGUSTUS

2013

PROCEEDINGS

SEMINAR NASIONAL TEKNIK INDUSTRI

TEMA:

*"Peluang dan Tantangan Globalisasi, Industrialisasi
dan kelestarian lingkungan untuk meningkatkan
kualitas hidup Masyarakat"*



Editor :

Dr. M. Sayuti, ST.,M.Sc.Eng

Fatimah, ST.,MT

Ir. Amri, MT

Diana Khairani Sofyan, ST.,MT

Syarifuddin, ST., MT

Jurusan Teknik Industri

Fakultas Teknik, Universitas Malikussaleh

Lhokseumawe-Aceh

PROCEEDINGS

DAFTAR ISI

Kata Pengantar	ii
Susunan Panitia	iv
Sambutan Ketua Panitia	v
Sambutan Ketua Jurusan	vi
Sambutan Dekan Fakultas Teknik	vii
Sambutan Rektor Universitas Malikussaleh	viii
Pembicara Utama	ix
Daftar Isi	x

ERGONOMI

Yunizurwan	1-8
ANALISIS PENGARUH WAKTU KERJA/SHIFT TERHADAP BEBAN KERJA MENTAL PENGAWAS PADA BAGIAN PRODUKSI PT.BSI PADANG DENGAN METODE NASA TLX	
Siti Rohmah	9-16
ANALISIS TINGKAT RISIKO ERGONOMI PADA PEKERJA KONVEKSI DENGAN METODA <i>JOB STRAIN INDEX</i> (JSI)	
Suanda Sufi dan Sayuti, M	17-22
ANALISA KERJA OPERATOR MESIN LAS DENGAN METODE WORK SAMPLING DI PT.X	
Roberta Zulfhi Surya dan Hari Purnomo	23-28
APLIKASI ERGONOMI DAN <i>VALUE ENGINEERING</i> DALAM PERANCANGAN <i>BILLBOARD</i> (STUDI KASUS PADA <i>BILLBOARD</i> KAMPANYE CALON BUPATI SLEMAN, YOGYAKARTA 2010)	
Fatimah, Amri dan Veronika Br Sitepu	29-36
APLIKASI SKOR REBA DAN ANAVA PADA PENENTUAN PENGARUH BEBAN KERJA FISIK TERHADAP KELELAHAN OTOT	
Anizar, Farida Ariani dan Idhar Yahya	37-45
KAJIAN ERGONOMI FASILITAS KERJA DI STASIUN PENCETAKAN KERUPUK IKAN	

MANAJEMEN PERAWATAN

- Dewi Mulyati, Bustami Syam dan Nazaruddin** 99-106
DAMPAK PENERAPAN TOTAL PRODUCTIVE MAINTENANCE TERHADAP
KINERJA PADA PT.BAHARI DWIKENCANA LESTARI (STUDI KASUS: PT.
BAHARI DWIKENCANA LESTARI)
- Syarifuddin** 107-114
PERANCANGAN SISTEM PEMELIHARAAN MESIN DENGAN
PENDEKATAN *RELIABILITY ENGINEERING*
- Budhi Santri Kusuma** 115-119
ANALISA MINIMALISASI KERUSAKAN ALAT DI PT X DENGAN
MEMAKSIMALKAN 8 PILAR PERAWATAN PRODUKTIF
- Amri, Fatimah dan Afzal** 120-127
PENGARUH JENIS DAN CAIRAN RADIATOR TERHADAP WAKTU
SERVICE

CAD/CAM/CAE, AUTOMASI, KONTROL DAN SIMULASI

- Suhendrianto** 128-134
PERANCANGAN *CNC TURNING MACHINE* DENGAN METODE
FUNCTIONAL MODELLING
- Teuku Syibran Malasy dan Sayuti, M** 135-141
PENDEKATAN METODE SIMULASI MONTE CARLO UNTUK
MENENTUKAN JUMLAH PENGGUNAAN JAM KERJA EXCAVATOR
CAT 320 A PADA PT. X
- Dody Sufriyadi dan Sayuti, M** 142-148
APLIKASI SIMULASI ANTRIAN TERHADAP MESIN BUFFING CR 50
UNTUK MENGETAHUI EFESIENSI MESIN PADA CV. X
- Aceng Komarudin Mutaqin, Abdul Kudus dan Fanny Tania Safitri** 149-156
PENDUGAAN PARAMETER DISTRIBUSI LOG-LOGISTIK UNTUK DATA
YANG MENGANDUNG PENGAMATAN TIDAK TERDETEKSI
- 157-163

PENDUGAAN PARAMETER DISTRIBUSI LOG-LOGISTIK UNTUK DATA YANG MENGANDUNG PENGAMATAN TIDAK TERDETEKSI

Aceng Komarudin Mutaqin, Abdul Kudus dan Fanny Tania Safitri

Program Studi Statistika,
Fakultas Matematika & Ilmu Pengetahuan Alam Universitas Islam Bandung
email: aceng.k.mutaqin@unisba.ac.id

Abstrak

Makalah ini membahas pendugaan parameter distribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi. Metode pendugaan yang digunakan adalah algoritme EM. Metode ini dibandingkan dengan metode yang telah ada, yaitu metode substitusi dengan menggunakan simulasi Monte Carlo. Hasilnya menunjukkan bahwa secara umum metode yang diusulkan lebih baik dibandingkan dengan metode substitusi ketika variansi dari datanya adalah kecil. Sebagai bahan aplikasi digunakan data sekunder mengenai konsentrasi tembaga dalam air tanah di lembah San Joaquin, California.

Kata Kunci: distribusi log-logistik, algoritme EM, simulasi Monte Carlo, pengamatan tidak terdeteksi.

Pendahuluan

Adanya pengamatan yang tidak terdeteksi seringkali dihadapi dalam menganalisis data lingkungan, misalnya konsentrasi suatu zat nilainya terlalu kecil untuk dapat diukur, sehingga dicatat sebagai n.d. (*nondetect*). Masalah pendugaan parameter dan penarikan kesimpulan berdasarkan data yang mengandung pengamatan tidak terdeteksi sangat menantang untuk diteliti, karena kita tidak dapat dengan mudah membuang data tersebut tanpa mengakibatkan adanya masalah bias dalam pendugaan parameter.

Biasanya pendekatan yang digunakan untuk menduga 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 batas deteksi (BD), yakni konsentrasi terkecil dari zat kimia yang dapat dibedakan dari konsentrasi nol. Biasanya praktisi mengganti n.d. dengan $BD/2$, $BD/\sqrt{2}$, atau nilai nol, kemudian menggunakan metode statistika untuk data lengkap. Cara ini mungkin tidak masalah untuk tujuan praktis jika persentase pengamatan n.d. sangat kecil, namun akan menjadi masalah jika persentasenya cukup besar. Tidak ada alasan rasional mengganti pengamatan n.d. dengan cara di atas. Lebih parah lagi, hasil penarikan kesimpulannya sangat tidak akurat.

Pendekatan parametrik mengasumsikan data mengikuti distribusi tertentu. Jika ukuran sampelnya cukup dan distribusi yang dipilih hampir sama dengan distribusi data aktual, maka pendugaan kemungkinan maksimum bersifat efisien. Tetapi mungkin saja syarat tersebut tidak terpenuhi, sehingga penduganya tidak lagi bersifat efisien. Pendekatan parametrik mempunyai kinerja yang buruk untuk data sampel berukuran antara 25 sampai 50 [2].

Pendekatan nonparametrik yang banyak dikembangkan dalam kajian analisis survival telah diadopsi untuk memecahkan masalah pendugaan parameter populasi yang data sampelnya mengandung pengamatan tidak terdeteksi. Pendekatan ini tidak mengasumsikan apapun mengenai bentuk distribusi datanya dan kinerja penduga nonparametrik ini cukup baik untuk ukuran sampel kecil ($n < 50$) dan persentase pengamatan tidak terdeteksinya dalam tingkat yang sedang [1].

Makalah ini bertujuan membahas sebuah metode parametrik untuk menduga parameter distribusi log-logistik berdasarkan data sampel yang mengandung pengamatan tidak terdeteksi. Metode tersebut adalah metode kemungkinan maksimum yang diintegrasikan dengan algoritme EM. Distribusi log-logistik dipertimbangkan dalam makalah ini karena merupakan salah satu distribusi yang dapat diaplikasikan pada data mutu lingkungan [4]. Kinerja dari metode yang diusulkan dibandingkan dengan metode substitusi melalui simulasi Monte Carlo. sebagai bahan aplikasi dari metode yang diusulkan adalah data mengenai konsentrasi tembaga dalam air tanah di lembah San Joaquin, California.

Distribusi Log-logistik

Salah satu distribusi yang dapat diaplikasikan untuk data mutu lingkungan adalah distribusi log-logistik diperumum [4]. Distribusi log-logistik adalah distribusi khusus dari distribusi log-logistik diperumum, dimana bentuk fungsi densitasnya adalah

$$g(x; \alpha, \beta) = \frac{\alpha}{x} \left[\frac{e^{\beta x^{\alpha}}}{(1 + e^{\beta x^{\alpha}})^2} \right]; x > 0, \quad (1)$$

dimana $\alpha > 0$ adalah parameter skala, dan $-\infty < \beta < \infty$ adalah parameter lokasi. Momen ke- k untuk distribusi log-logistik [3] di atas adalah

$$E[X^k] = e^{-k\beta/\alpha} \Gamma\left(1 + \frac{k}{\alpha}\right) \Gamma\left(1 - \frac{k}{\alpha}\right); -\alpha < k < \alpha,$$

Sedangkan fungsi distribusinya adalah

$$G(x; \alpha, \beta) = \left(\frac{e^{\beta x^{\alpha}}}{1 + e^{\beta x^{\alpha}}} \right); x > 0. \quad (2)$$

Pendugaan Parameter Distribusi Log-logistik Melalui Algoritme EM

Asumsikan bahwa data mutu lingkungan berasal dari populasi yang berdistribusi log-logistik. Misalkan x_i menyatakan pengamatan terdeteksi ke- i , dengan $i = 1, 2, \dots, n$; p menyatakan banyaknya jenis alat ukur dengan BD berbeda-beda, dan t_j menyatakan banyaknya pengamatan tidak terdeteksi untuk BD_j , dengan $j = 1, \dots, p$, maka fungsi kemungkinannya adalah

$$L(\alpha, \beta) = \prod_{i=1}^n g(x_i; \alpha, \beta) \left\{ \prod_{j=1}^p \left(\prod_{m=1}^{t_j} G(BD_j; \alpha, \beta) \right) \right\}.$$

Dengan demikian fungsi log-kemungkinannya adalah

$$l_1(\alpha, \beta) = n \ln(\alpha) + n\beta + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - 2 \sum_{i=1}^n \ln(1 + e^{\beta} x_i^{\alpha}) \\ + \beta \sum_{j=1}^p t_j + \alpha \sum_{j=1}^p t_j \ln(BD_j) - \sum_{j=1}^p t_j \ln(1 + e^{\beta} BD_j^{\alpha}).$$

Berdasarkan fungsi log-kemungkinan di atas, untuk mengintegrasikan pendugaan kemungkinan maksimum dengan algoritme EM, perlu dihitung nilai ekspektasi dari $\ln(X_j)|X_j < BD_j$ dan $\ln(1 + e^{\beta} X_j^{\alpha})|X_j < BD_j$, untuk $j = 1, \dots, p$. Dapat ditunjukkan bahwa untuk $j = 1, \dots, p$, nilai ekspektasi

$$E[\ln(X_j)|X_j < BD_j] = -\frac{\beta}{\alpha} + \frac{\ln(G(BD_j))}{\alpha} + \frac{\left[\frac{1}{G(BD_j)} - 1\right] \ln(1 - G(BD_j))}{\alpha}, \quad (3)$$

dan

$$E[\ln(1 + e^{\beta} X_j^{\alpha})|X_j < BD_j] = \ln\left(1 + \frac{G(BD_j)}{1 - G(BD_j)}\right) + \frac{\ln(1 - G(BD_j))}{G(BD_j)} + 1. \quad (4)$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln(x_j)$ dan $\ln(1 + e^{\beta} x_j^{\alpha})$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (3) dan (4). Misalkan $\alpha^{(r)}$ dan $\beta^{(r)}$ adalah taksiran parameter α dan β pada iterasi ke- r , tahap-M adalah memaksimumkan fungsi kemungkinan berikut untuk memperoleh taksiran parameter α dan β pada iterasi ke- $r + 1$

$$l_2(\alpha, \beta) = n \ln(\alpha) + n\beta + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - 2 \sum_{i=1}^n \ln(1 + e^{\beta} x_i^{\alpha}) \\ + \ln(\alpha) \sum_{j=1}^p t_j + \beta \sum_{j=1}^p t_j + (\alpha - 1) \sum_{j=1}^p t_j E_j^1 - 2 \sum_{j=1}^p t_j E_j^2, \quad (5)$$

dimana E_j^1 dan E_j^2 masing-masing adalah ekspektasi yang ada pada Persamaan (3) dan (4), dengan α dan β diganti oleh $\alpha^{(r)}$ dan $\beta^{(r)}$. Solusi pada tahap-M tidak dapat diperoleh secara analitik, sehingga perlu dicari menggunakan metode numerik, dalam hal ini metode Newton-Raphson. Turunan pertama dari fungsi log-kemungkinan pada Persamaan (5) terhadap parameter α dan β masing-masing adalah :

$$\frac{\partial l_2}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln(x_i) - 2 \sum_{i=1}^n \frac{e^{\beta} x_i^{\alpha} \ln(x_i)}{1 + e^{\beta} x_i^{\alpha}} + \frac{1}{\alpha} \sum_{j=1}^p t_j + \sum_{j=1}^p t_j E_j^1 \quad (6)$$

dan

$$\frac{\partial E[l(\alpha, \beta)]}{\partial \beta} = n - 2 \sum_{i=1}^n \frac{e^{\beta} x_i^{\alpha}}{1 + e^{\beta} x_i^{\alpha}} + \sum_{j=1}^p t_j. \quad (7)$$

Turunan kedua dari fungsi log-kemungkinan pada Persamaan (5) terhadap parameter α dan β masing-masing adalah:

$$\frac{\partial^2 l_2}{\partial \alpha^2} = -\frac{n}{\alpha^2} - 2 \sum_{i=1}^n \frac{e^{\beta x_i^\alpha} [\ln(x_i)]^2}{(1 + e^{\beta x_i^\alpha})^2} - \frac{1}{\alpha^2} \sum_{j=1}^p t_j, \quad (8)$$

$$\frac{\partial^2 l_2}{\partial \beta^2} = -2 \sum_{i=1}^n \frac{e^{\beta x_i^\alpha}}{(1 + e^{\beta x_i^\alpha})^2}, \quad (9)$$

dan

$$\frac{\partial^2 l_2}{\partial \alpha \partial \beta} = -2 \sum_{i=1}^n \frac{e^{\beta x_i^\alpha} \ln(x_i)}{(1 + e^{\beta x_i^\alpha})^2}, \quad (10)$$

Berdasarkan turunan pertama dan kedua dari fungsi log-kemungkinan di atas, dapat diperoleh penduga parameter distribusi log-logistik.

Pendugaan kemungkinan maksimum dengan algoritme EM memerlukan nilai awal taksiran parameter agar proses iterasi bisa berjalan, nilai awal yang digunakan adalah nilai taksiran parameter α dan β menggunakan metode *percentile matching* [3] berdasarkan data yang terobservasi, yaitu:

$$\alpha^{(0)} = \frac{2 \ln(3)}{\ln(w) - \ln(v)}, \quad (11)$$

dan

$$\beta^{(0)} = -\ln(3) \frac{\ln(w) + \ln(v)}{\ln(w) - \ln(v)}, \quad (12)$$

dimana v dan w masing-masing menyatakan persentil ke-25 dan ke-75 dari data yang terobservasi.

Studi Simulasi

Dalam bagian ini akan dilakukan perbandingan metode yang diusulkan, yaitu metode pendugaan kemungkinan maksimum melalui algoritme EM dengan metode substitusi yang sering digunakan oleh praktisi, yaitu $BD/2$. Perbandingannya akan dilakukan dengan menggunakan simulasi Monte Carlo untuk melihat kinerja dari kedua metode di atas. Ukuran yang dijadikan perbandingan adalah bias relatif dan rata-rata jumlah kuadrat kekeliruan.

Data sensor dengan BD ganda akan dibangkitkan dengan skema sampling sebagaimana dalam [5]. Data dibangkitkan melalui 3 tahap. Tahap 1, peubah A berukuran N ($n + \sum_{j=1}^p t_j$) disimulasikan dari distribusi log-logistik tertentu, peubah B dengan ukuran yang sama disimulasikan dari distribusi seragam dalam $(0,1)$. A dan B digabungkan untuk membentuk matriks C berukuran $N \times 2$, dimana kolom pertama berisikan A , dan kolom kedua B . Tahap 2, matriks C diurut berdasarkan kolom kedua, kemudian kolom pertama dibagi 6 subgrup sama banyak. Tahap 3, BD_i untuk

setiap subgrup i , $i = 1, 2, \dots, 6$, dipilih sedemikian sehingga peluang $P(A_i < BD_i) = p_i$, yang menunjukkan rata-rata ada data tersensor sebanyak $100p_i$ persen dari data dalam setiap subgrup yang disimulasikan.

Tabel 1. Bias relatif dalam menaksir parameter α

Kasus	Ukuran Sampel	Persentase Sensor	Variansi	Metode Penaksiran	
				EM	Substitusi
1	Kecil	Rendah	Rendah	-0,0357	-0,1990
2	Kecil	Sedang	Rendah	-0,3046	NaN*
4	Sedang	Rendah	Rendah	-0,0624	-0,2334
5	Sedang	Sedang	Rendah	-0,4394	NaN
7	Besar	Rendah	Rendah	-0,0665	-0,2354
8	Besar	Sedang	Rendah	-0,4440	NaN
10	Kecil	Rendah	Sedang	-0,0371	-0,0406
11	Kecil	Sedang	Sedang	-0,3036	NaN
13	Sedang	Rendah	Sedang	-0,0634	-0,0613
14	Sedang	Sedang	Sedang	-0,4400	-0,1276
16	Besar	Rendah	Sedang	-0,0662	-0,0651
17	Besar	Sedang	Sedang	-0,4447	-0,0757
19	Kecil	Rendah	Tinggi	-0,0380	0,0633
20	Kecil	Sedang	Tinggi	-0,3053	0,1931
22	Sedang	Rendah	Tinggi	-0,0621	0,0650
23	Sedang	Sedang	Tinggi	-0,4393	0,2056
25	Besar	Rendah	Tinggi	-0,0660	0,0597
26	Besar	Sedang	Tinggi	-0,4437	0,1994

*NaN: not a number

Tabel 2. Rata-rata jumlah kuadrat kekeliruan dalam menaksir parameter α

Kasus	Ukuran Sampel	Persentase Sensor	Variansi	Metode Penaksiran	
				EM	Substitusi
1	Kecil	Rendah	Rendah	0,2255	1,2218
2	Kecil	Sedang	Rendah	2,2706	NaN*
4	Sedang	Rendah	Rendah	0,1555	1,2705
5	Sedang	Sedang	Rendah	4,3216	NaN
7	Besar	Rendah	Rendah	0,1324	1,2394
8	Besar	Sedang	Rendah	4,3933	NaN
10	Kecil	Rendah	Sedang	0,0580	0,0423
11	Kecil	Sedang	Sedang	0,5661	NaN
13	Sedang	Rendah	Sedang	0,0401	0,0318
14	Sedang	Sedang	Sedang	1,0946	0,8096
16	Besar	Rendah	Sedang	0,0329	0,0288
17	Besar	Sedang	Sedang	1,1131	0,1574
19	Kecil	Rendah	Tinggi	0,0122	0,0153
20	Kecil	Sedang	Tinggi	0,1238	0,0624
22	Sedang	Rendah	Tinggi	0,0085	0,0090
23	Sedang	Sedang	Tinggi	0,2308	0,0553
25	Besar	Rendah	Tinggi	0,0070	0,0061
26	Besar	Sedang	Tinggi	0,2345	0,0498

*NaN: not a number

Untuk melihat pengaruh ukuran sampel, persentase data tersensor, dan variabilitas dalam data, beberapa kasus akan dicobakan dalam simulasi. Ukuran sampel yang akan dicobakan adalah kecil ($N = 90$), sedang ($N = 240$) dan besar ($N = 480$). Persentase data tersensor yang digunakan adalah rendah ($p_1 = 0,3; p_2 = 0,25; p_3 = 0,2; p_4 = 0,15; p_5 = 0,1; p_6 = 0,05$), sedang ($p_1 = 0,5; p_2 = 0,45; p_3 = 0,4; p_4 = 0,35; p_5 = 0,3; p_6 = 0,25$), dan tinggi ($p_1 = 0,75; p_2 = 0,7; p_3 = 0,65; p_4 = 0,6; p_5 =$

0,55; $p_6 = 0,5$). Sedangkan variabilitas data yang akan dicobakan adalah rendah ($\alpha = 4,7110$; $\beta = -6,3344$), sedang ($\alpha = 2,3679$; $\beta = -3,1982$), dan tinggi ($\alpha = 1,0892$; $\beta = -1,4404$). Demikian akan ada 27 kasus yang dicobakan. Simulasi akan dilakukan sebanyak 5.000 kali. Hasil simulasi disajikan dalam Tabel 1 sampai Tabel 4.

Tabel 3. Bias relatif dalam menaksir parameter β

Kasus	Ukuran Sampel	Persentase Sensor	Variansi	Metode Penaksiran	
				EM	Substitusi
1	Kecil	Rendah	Rendah	-0,0546	-0,2163
2	Kecil	Sedang	Rendah	-0,4214	NaN*
4	Sedang	Rendah	Rendah	-0,0858	-0,2548
5	Sedang	Sedang	Rendah	-0,5906	NaN
7	Besar	Rendah	Rendah	-0,0891	-0,2566
8	Besar	Sedang	Rendah	-0,5943	NaN
10	Kecil	Rendah	Sedang	-0,0742	-0,0560
11	Kecil	Sedang	Sedang	-0,5349	NaN
13	Sedang	Rendah	Sedang	-0,1084	-0,0801
14	Sedang	Sedang	Sedang	-0,7392	-0,1800
16	Besar	Rendah	Sedang	-0,1098	-0,0838
17	Besar	Sedang	Sedang	-0,7424	-0,1326
19	Kecil	Rendah	Tinggi	-0,1225	0,0718
20	Kecil	Sedang	Tinggi	-0,8166	0,2633
22	Sedang	Rendah	Tinggi	-0,1663	0,0703
23	Sedang	Sedang	Tinggi	-1,1029	0,2737
25	Besar	Rendah	Tinggi	-0,1657	0,0649
26	Besar	Sedang	Tinggi	-1,1055	0,2576

*NaN: not a number

Tabel 4. Rata-rata jumlah kuadrat kekeliruan dalam menaksir parameter β

Kasus	Ukuran Sampel	Persentase Sensor	Variansi	Metode Penaksiran	
				EM	Substitusi
1	Kecil	Rendah	Rendah	0,5159	2,4980
2	Kecil	Sedang	Rendah	7,5945	NaN*
4	Sedang	Rendah	Rendah	0,4324	2,7185
5	Sedang	Sedang	Rendah	14,0738	NaN
7	Besar	Rendah	Rendah	0,3874	2,6607
8	Besar	Sedang	Rendah	14,2065	NaN
10	Kecil	Rendah	Sedang	0,1861	0,1238
11	Kecil	Sedang	Sedang	3,0652	NaN
13	Sedang	Rendah	Sedang	0,1653	0,0959
14	Sedang	Sedang	Sedang	5,6139	1,4994
16	Besar	Rendah	Sedang	0,1449	0,0862
17	Besar	Sedang	Sedang	5,6488	0,3864
19	Kecil	Rendah	Tinggi	0,0827	0,0674
20	Kecil	Sedang	Tinggi	1,4484	0,2329
22	Sedang	Rendah	Tinggi	0,0762	0,0315
23	Sedang	Sedang	Tinggi	2,5331	0,1854
25	Besar	Rendah	Tinggi	0,0661	0,0191
26	Besar	Sedang	Tinggi	2,5400	0,1525

*NaN: not a number

Tabel 1 dan 2 masing-masing berisikan bias relatif dan rata-rata jumlah kuadrat kekeliruan untuk menaksir parameter α . Sedangkan Tabel 3 dan 4 masing-masing berisikan bias relatif dan rata-rata jumlah kuadrat kekeliruan untuk menaksir parameter β . Terlihat bahwa kasus 3, 6, 9, 12, 15, 18, 21, 24, dan 27 (kasus untuk persentase sensor tinggi) tidak ada dalam tabel-tabel di atas karena semua hasilnya

adalah NaN (*not a number*). Salah satu penyebabnya adalah matriks turunan keduanya singular. Hasil NaN juga didapatkan oleh metode substitusi pada kasus 2, 5, 8 (kasus semua ukuran sampel, persentase sensor sedang dan variansi rendah), dan 11 (ukuran sampel kecil, persentase sensor sedang dan variansi sedang). Dengan demikian untuk keempat kasus tersebut metode pendugaan kemungkinan maksimum melalui algoritme EM yang cocok digunakan untuk menduga parameter α dan β . Terlihat jelas bahwa metode pendugaan kemungkinan maksimum melalui algoritme EM lebih baik dibandingkan dengan metode substitusi dalam menduga parameter α dan β untuk kasus variansi data rendah, ukuran sampel kecil-sedang dan persentase sensor rendah-sedang. Metode yang diusulkan juga lebih baik dibandingkan dengan metode substitusi dalam menduga parameter α untuk kasus variansi data tinggi, persentase sensor rendah dan ukuran sampel kecil-sedang. Untuk kasus yang sama metode substitusi yang lebih baik dalam menduga parameter β . Metode substitusi lebih baik dibandingkan dengan metode yang diusulkan dalam menduga parameter α dan β untuk kasus variansi data sedang, ukuran sampel sedang-besar, dan persentase sensor rendah-sedang. Hasil yang sama juga ditunjukkan untuk kasus variansi data tinggi, persentase sensor sedang, dan semua ukuran sampel. Hasil yang tidak berbeda ditunjukkan pula untuk kasus variansi data tinggi, persentase sensor rendah, dan ukuran sampel besar. Untuk kasus ukuran sampel besar, persentase sensor rendah-tinggi, dan variansi data sedang-tinggi, metode substitusi lebih baik dibandingkan dengan metode yang diusulkan dalam menduga parameter α dan β .

Contoh Numerik

Data yang digunakan untuk mengaplikasikan metode pendugaan kemungkinan maksimum yang diintegrasikan dengan algoritme EM adalah data mengenai konsentrasi tembaga dalam air tanah di lembah San Joaquin, California [4].

Tabel 5. Konsentrasi* tembaga (Cu) dalam air tanah dalam satuan mikrogram per liter di lembah San Joaquin, California

2	4	2	<5	<15	8	3
2	<10	<10	17	<5	1	6
12	<1	3	23	4	15	3
2	1	<1	9	<4	3	4
1	<2	1	9	<5	3	5
<10	<2	1	3	<5	1	14
<10	1	3	3	4	6	4

*tanda < menunjukkan pengamatan tidak terdeteksi berdasarkan alat ukur yang mempunyai BD tertentu.

Diasumsikan bahwa data dalam Tabel 5 tersebut berasal dari populasi yang berdistribusi log-logistik. Hasil pendugaan parameter distribusi log-logistiknya adalah $\hat{\alpha} = 1,6292$ dan $\hat{\beta} = -0,8398$. Dengan menggunakan prinsip plug-in, rata-rata dari distribusi log-logistik tersebut adalah 3,4467. Jadi disimpulkan bahwa rata-rata konsentrasi tembaga dalam air tanah di lembah San Joaquin, California adalah 3,4467 mikrogram per liter.

Kesimpulan

Dalam makalah ini telah dirumuskan metode pendugaan kemungkinan maksimum yang terintegrasi dengan algoritme EM untuk menduga parameter distribusi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi. Hasil simulasi menunjukkan bahwa secara umum metode pendugaan kemungkinan maksimum melalui algoritme EM lebih baik dibandingkan dengan metode substitusi dalam menduga parameter distribusi log-logistik untuk kasus variansi data rendah, berbagai ukuran sampel dan persentase sensor rendah-sedang. Sedangkan metode substitusi secara umum lebih baik dibandingkan dengan metode yang diusulkan untuk kasus variansi data sedang dan tinggi. Walaupun demikian ada kasus-kasus tertentu yang menunjukkan tidak demikian.

Makalah ini membahas masalah pendugaan parameter distribusi log-logistik untuk data sampel yang mengandung pengamatan tidak terdeteksi. Masalah yang dibahas adalah untuk kasus data satu sampel. Untuk itu pada penelitian selanjutnya, tim peneliti berencana untuk merumuskan masalah yang terkait dengan data dua sampel atau lebih yang mengandung pengamatan tidak terdeteksi.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Lembaga Penelitian dan Pengabdian pada Masyarakat (LPPM) Unisba Bandung atas dukungan finansialnya. Penelitian ini dibiayai oleh LPPM Unisba Bandung dengan Nomor Kontrak: 273/LPPM-SP3/XII/2012.

Daftar Pustaka

- [1] R.O. Gilbert, *Statistical Method for Environmental Pollution Monitoring* (Wiley, New York 1987).
- [2] Gleit A., Estimation for small normal data sets with detection limits. *Environmental Science and Technology*, **19** (1985), 1201-1206.
- [3] S.A. Klugman, H.H. Panjer, dan G.E. Willmot, *Loss Models: From Data to Decisions*, (Wiley, New York 2004).
- [4] Warsono, *Analysis of Environmental Pollutant Data Using Generalized Log-logistic Distribution*, Dissertation at University of Alabama at Birmingham, 1996.
- [5] W. Zhong, R. Shukla, P. Succop, L. Levin, J. Welge, dan S. Sivaganesan, *Statistical Approaches to Analyze Censored Data with Multiple Detection Limits*. Disertasi Program Doctor of Philosophy University of Cincinnati, 2005.