LAPORAN TAHUN TERAKHIR

PENELITIAN FUNDAMENTAL

PENDEKATAN STATISTIK UNTUK PERBANDINGAN DUA SAMPEL YANG MENGANDUNG PENGAMATAN TIDAK TERDETEKSI DENGAN PENERAPAN PADA DATA LINGKUNGAN

Tahun ke 2 dari rencana 2 tahun

KETUA/ANGGOTA TIM

Dr. Aceng Komarudin Mutaqin, S.Si., MT., M.Si. (NIDN: 0428117401)

Siti Sunendiari, Dra., M.Si. (NIDN: 0422106101)

UNIVERSITAS ISLAM BANDUNG

November 2016

HALAMAN PENGESAHAN

Judul	: PENDEKATAN STATISTIK UNTUK PERBANDINGAN DUA SAMPEL YANG MENGANDUNG PENGAMATAN TIDAK TERDETEKSI DENGAN PENERAPAN PADA DATA LINGKUNGAN
Peneliti/Pelaksana	
Nama Lengkap	: Dr. ACENG KOMARUDIN MUTAQIN S.Si., M.T., M.Si.
Perguruan Tinggi	: Universitas Islam Bandung
NIDN	: 0428117401
Jabatan Fungsional	: Lektor
Program Studi	: Statistika
Nomor HP	: 08179289987
Alamat surel (e-mail)	: aceng.k.mutaqin@gmail.com
Anggota (1)	
Nama Lengkap	: Dra SITI SUNENDIARI M.Si
NIDN	: 0422106101
Perguruan Tinggi	: Universitas Islam Bandung
Institusi Mitra (jika ada)	
Nama Institusi Mitra	: -
Alamat	: -
Penanggung Jawab	: -
Tahun Pelaksanaan	: Tahun ke 2 dari rencana 2 tahun
Biaya Tahun Berjalan	: Rp 50.000.000,00
Biaya Keseluruhan	: Rp 100.000.000,00

17 6191

Mengetahui,
Dekan FMIPA Unisba

(Dr. Suwanda, MS.)
NIP/NIK D.86.0.045

Bandung, 14 - 11 - 2016
Ketua,

(Dr. ACENG KOMARUDIN MUTAQIN S.Si., M.T., M.Si.)
NIP/NIK D.01.0.347

Menyetujui,
Ketua LPPM Unisba

(Prof. Dr. H. Edi Setiadi, SH., MH.)
NIP/NIK 195911101987031002

RINGKASAN

Penelitian ini membahas pendekatan statistik untuk membandingkan dua sampel saling bebas berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi. Untuk kasus data tersebut, penelitian awal telah dilakukan untuk kasus satu sampel dalam menaksir parameter dari distribusi log-logistik menggunakan metode pendugaan kemungkinan maksimum melalui algoritme EM (Ekspektasi-Maksimisasi).
Penulis telah melakukan penelitian untuk membandingkan dua sampel saling bebas berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi menggunakan uji permutasi dengan menguji kesamaan dua median populasinya. Program MATLAB telah dibuat untuk pengujian di atas. Hasil simulasi menunjukkan bahwa (1) untuk data yang mengandung batas deteksi tunggal, uji permutasi untuk menguji kesamaan dua median dari dua populasi berdistribusi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi rendah dan perbedaan besar koefisien variasi; (2) untuk data yang mengandung batas deteksi ganda, uji permutasi untuk menguji kesamaan dua median dari dua populasi berdistribusi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi kedua sampel berbeda dan perbedaan besar koefisien variasi.
Dalam penelitian ini telah dihasilkan metode lain untuk membandingkan dua sampel berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi menggunakan informasi fungsi kemungkinan. Diasumsikan bahwa dua sampel saling bebas berukuran n_1 dan n_2 berasal dari dua populasi log-logistik $LLD(\theta_1, \gamma_1)$ dan $LLD(\theta_2, \gamma_2)$. Penaksiran dengan menggunakan algoritme EM digunakan untuk menaksir parameter dalam empat kasus: $H_1(\theta_1 = \theta_2, \gamma_1 = \gamma_2)$, $H_2(\theta_1 \neq \theta_2, \gamma_1 = \gamma_2)$, $H_3(\theta_1 = \theta_2, \gamma_1 \neq \gamma_2)$, dan $H_1(\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2)$. Suatu pedoman digambarkan untuk menguji kesamaan dua median ($H_0: \theta_1 = \theta_2$ lawan $H_r: \theta_1 \neq \theta_2$). Dua prosedur diusulkan untuk uji di atas, tergantung apakah koefisien variasinya sama: ($H_0: \gamma_1 = \gamma_2$) atau tidak ($H_0: \gamma_1 \neq \gamma_2$). Uji chi-square asimtotik digunakan dalam uji di atas. Hasil simulasi Monte Carlo menunjukkan bahwa metode yang dibangun memiliki kinerja yang bagus terutama untuk ukuran sampel yang besar. Dalam penelitian ini diberikan suatu contoh kasus menggunakan data emisi gas buang kendaraan bermotor.

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

PRAKATA

Segala puji hanya milik Alloh SWT. yang telah memberikan kekuatan kepada tim peneliti sehingga dapat menyelesaikan Laporan Tahun Terakhir Hibah Fundamental ini.

Penulis mengucapkan terima kasih kepada DIKTI yang telah memberikan bantuan berupa dukungan finansialnya. Penulis juga mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian pada Masyarakat (LPPM) Universitas Islam Bandung yang telah menjadi fasilitator dalam kegiatan penelitian ini.

Ucapan terima kasih juga disampaikan kepada Ketua Program Studi Statistika dan Dekan Fakultas MIPA Universitas Islam Bandung selaku pimpinan lembaga dimana kami bekerja atas dukungan morilnya.

Terakhir penulis ucapkan terima kasih kepada semua pihak yang belum disebutkan di atas, baik langsung ataupun tidak langsung berkontribusi atas terlaksananya penelitian ini.

Semoga semua kebaikan yang telah ditebarkan oleh semua pihak mendapatkan balasan dengan yang lebih baik dan lebih banyak dari yang Maha Kuasa.

Bandung, 30 November 2016

Ketua Tim Peneliti

Dr. Aceng Komarudin Mutaqin, MT., M.Si.

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB 1. PENDAHULUAN

Data lingkungan seringkali memuat nilai-nilai pengamatan yang berada di bawah batas deteksi, sehingga nilai pengamatan sebenarnya tidak terdeteksi atau teramati. Masalah-masalah statistik yang berkaitan dengan data yang mengandung pengamatan tidak terdeteksi sangat menantang untuk diteliti.

Pendekatan statistik yang biasa 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 nilai yang tergantung pada batas deteksi (BD) alat ukur. Biasanya praktisi menggantinya dengan nilai nol, BD, atau $BD/2$. Tidak ada alasan rasional mengganti pengamatan tidak terdeteksi dengan cara substitusi. Pendekatan parametrik mengasumsikan data mengikuti distribusi tertentu. Gleit (1985) dan Shumway dkk. (2002) menunjukkan bahwa pendekatan parametrik mempunyai kinerja yang buruk untuk data sampel berukuran antara 25 sampai 50. Dengan mengasumsikan data mengikuti distribusi log-logistik, Mutaqin dkk. (2013) menunjukkan bahwa metode pendugaan kemungkinan maksimum melalui algoritme EM (Ekspektasi-Maksimisasi) mempunyai kinerja yang bagus dibandingkan dengan metode substitusi ketika variansi datanya kecil. Pendekatan nonparametrik dalam kajian analisis survival telah diadopsi untuk memecahkan masalah yang dikemukakan di atas. Pendekatan ini cukup baik untuk ukuran sampel kecil (n<50) dan persentase pengamatan tidak terdeteksinya dalam tingkat yang sedang (Gilbert, 1987).

Nilai pengamatan yang tidak terdeteksi menjadi suatu masalah yang sulit ketika tujuannya adalah membandingkan dua populasi yang berbeda. Secara umum ada dua pendekatan yang diusulkan untuk permasalahan tersebut, yaitu pendekatan parametrik dan nonparametrik. Untuk dua data sampel dari dua populasi yang berbeda mengikuti distribusi lognormal, Stoline (1993) mengusulkan menggunakan uji kesamaan dua median untuk membandingkan dua populasi ketika data mengandung pengamatan tidak terdeteksi. Sementara itu, Zhong dkk. (2005) menggunakan informasi fungsi kemungkinan untuk pengujiannya. Untuk kasus yang sama, uji standar seperti uji T seringkali digunakan oleh para peneliti (Zhong dkk., 2005).

Selain distribusi lognormal, distribusi lain yang bisa digunakan untuk memodelkan data lingkungan adalah distribusi log-logistik (Warsono, 1996). Mutaqin dan Kudus (2014) membahas uji permutasi untuk membandingkan dua populasi berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi dengan jalan membandingkan kedua median populasinya. Hasil penelitian Mutaqin (2015) menunjukkan bahwa untuk data yang mengandung batas deteksi tunggal, uji permutasi tersebut akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi rendah dan perbedaan besar koefisien variasinya. Sedangkan untuk data yang mengandung batas deteksi ganda, uji permutasi tersebut akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi kedua sampel berbeda dan perbedaan besar koefisien variasi. Dalam penelitian ini akan diusulkan metode lain yang sifatnya parametrik yang menggunakan informasi fungsi kemungkinan (Stoline, 1993).

BAB 2. TINJAUAN PUSTAKA

2.1. Pendahuluan

Secara umum ada dua pendekatan yang diusulkan untuk masalah dua sampel yang mengandung pengamatan tidak terdeteksi, yaitu pendekatan parametrik dan nonparametrik. Untuk dua data sampel dari dua populasi yang berbeda mengikuti distribusi lognormal, Stoline (1993) mengusulkan menggunakan uji kesamaan dua median yang bersifat parametrik. Sementara itu, Zhong dkk. (2005) menggunakan informasi fungsi kemungkinan yang bersifat parametrik dan uji permutasi yang bersifat nonparametrik.

Selain distribusi lognormal, distribusi lain yang bisa digunakan untuk memodelkan data lingkungan adalah distribusi log-logistik (Warsono, 1996). Mutaqin dan Kudus (2014) membahas uji permutasi yang bersifat nonparametrik untuk membandingkan dua populasi berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi dengan jalan membandingkan kedua median populasinya. Hasil penelitian Mutaqin (2015) menunjukkan bahwa secara umum uji permutasi akan baik digunakan untuk kasus ukuran sampel besar, dan perbedaan besar koefisien variasinya.

2.2. Distribusi Log-Logistik

Distribusi log-logistik adalah distribusi khusus dari distribusi log-logistik diperumum yang bentuk fungsi densitasnya adalah

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

dimana $\alpha>0$ adalah parameter skala, dan $-\infty<\beta<\infty$ adalah parameter lokasi. Momen ke-k untuk distribusi log-logistik (Klugman dkk., 2004) 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

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

2.3. Penaksiran Parameter Distribusi Log-logistik Melalui Algoritme EM

Teori yang dibahas dalam bagian ini semuanya merupakan hasil dari Mutaqin dkk. (2013). Asumsikan bahwa data lingkungan berasal dari populasi yang berdistribusi log-logistik. Misalkan x_i menyatakan pengamatan terdeteksi ke-i, dengan $i = 1, 2, \cdots, n$; p menyatakan banyaknya jenis alat ukur dengan BD berbeda-beda, dan t_j menyatakan banyaknya pengamatan tidak terdeteksi untuk BD_j, dengan $j = 1, \cdots, p$.Untuk mengintegrasikan penaksiran 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 $j = 1, \cdots, p$. Dapat ditunjukkan bahwa untuk $j = 1, \cdots, p$, nilai ekspektasi

$$E\left[\ln(X_j)|X_j < BD_j\right] = -\frac{\beta}{\alpha} + \frac{\ln\left(G(BD_j)\right)}{\alpha} + \frac{\left[\frac{1}{G(BD_j)} - 1\right]\ln\left(1 - G(BD_j)\right)}{\alpha}, \quad (2.1)$$

dan

$$E\left[\ln(1 + e^{\beta} X_j{}^{\alpha})|X_j < BD_j\right] = \ln\left(1 + \frac{G(BD_j)}{1 - G(BD_j)}\right) + \frac{\ln\left(1 - G(BD_j)\right)}{G(BD_j)} + 1. \quad (2.2)$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln\left(x_j\right)$ dan $\ln\left(1 + e^{\beta} x_j{}^{\alpha}\right)$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (2.1) dan (2.2). 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-$r + 1$

$$\begin{aligned} l_2(\alpha, \beta) = {} & n\ln(\alpha) + n\beta + (\alpha - 1)\sum_{i=1}^{n} \ln(x_i) - 2\sum_{i=1}^{n} \ln\left(1 + e^{\beta} x_i{}^{\alpha}\right) \\ & + \ln(\alpha)\sum_{j=1}^{p} t_j + \beta\sum_{j=1}^{p} t_j + (\alpha - 1)\sum_{j=1}^{p} t_j E_j^1 - 2\sum_{j=1}^{p} t_j E_j^2, \end{aligned} \quad (2.3)$$

dimana E_j^1 dan E_j^2 masing-masing adalah ekspektasi yang ada pada Persamaan (2.1) dan (2.2), 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, salah satunya adalah metode Newton-Raphson. Turunan pertama dari fungsi log-kemungkinan pada Persamaan (2.3) terhadap parameter α dan β masing-masing adalah

$$\frac{\partial l_2}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \ln(x_i) - 2\sum_{i=1}^{n} \frac{e^{\beta} {x_i}^{\alpha} \ln(x_i)}{1 + e^{\beta} {x_i}^{\alpha}} + \frac{1}{\alpha}\sum_{j=1}^{p} t_j + \sum_{j=1}^{p} t_j E_j^1 \tag{2.4}$$

dan

$$\frac{\partial E[l(\alpha,\beta)]}{\partial \beta} = n - 2\sum_{i=1}^{n} \frac{e^{\beta} {x_i}^{\alpha}}{1 + e^{\beta} {x_i}^{\alpha}} + \sum_{j=1}^{p} t_j. \tag{2.5}$$

Turunan kedua dari fungsi log-kemungkinan pada Persamaan (2.3) terhadap parameter α dan β masing-masing adalah

$$\frac{\partial^2 l_2}{\partial \alpha^2} = -\frac{n}{\alpha^2} - 2\sum_{i=1}^{n} \frac{e^{\beta} {x_i}^{\alpha} [\ln(x_i)]^2}{(1 + e^{\beta} {x_i}^{\alpha})^2} - \frac{1}{\alpha^2}\sum_{j=1}^{p} t_j, \tag{2.6}$$

$$\frac{\partial^2 l_2}{\partial \beta^2} = -2\sum_{i=1}^{n} \frac{e^{\beta} {x_i}^{\alpha}}{(1 + e^{\beta} {x_i}^{\alpha})^2}, \tag{2.7}$$

dan

$$\frac{\partial^2 l_2}{\partial \alpha \partial \beta} = -2\sum_{i=1}^{n} \frac{e^{\beta} {x_i}^{\alpha} \ln(x_i)}{(1 + e^{\beta} {x_i}^{\alpha})^2}, \tag{2.8}$$

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

2.4. Uji Permutasi untuk Kesamaan Dua Median dari Distribusi Log-logistik

Distribusi log-logistik yang ada pada Subbab 2.1 akan diparameterisasi ulang melalui hubungan $\gamma = \alpha$, dan $\theta = e^{-\beta/\alpha}$. Melalui parameterisasi ulang tersebut, dapat ditunjukkan dengan mudah bahwa fungsi densitas dari distribusi log-logistiknya menjadi

$$f(x;\theta,\gamma) = \frac{\gamma(x/\theta)^{\gamma}}{x[1 + (x/\theta)^{\gamma}]^2}; x > 0, \tag{2.9}$$

dimana $\gamma > 0$ adalah parameter bentuk, dan $\theta > 0$ adalah parameter skala. Momen ke-k untuk distribusi log-logistik (Klugman dkk., 2008) di atas adalah

$$E[X^k] = \theta^k \Gamma\left(1 + \frac{k}{\gamma}\right)\Gamma\left(1 - \frac{k}{\gamma}\right); \ -\gamma < k < \gamma, \tag{2.10}$$

sedangkan fungsi distribusinya adalah

$$F(x;\theta,\gamma)=\frac{(x/\theta)^{\gamma}}{1+(x/\theta)^{\gamma}};x>0.$$

Misalkan M adalah median dari distribusi log-logistik di atas, maka $M=\theta$, yang hanya tergantung pada parameter θ. Sedangkan koefisien variasi dari distribusi log-logistik di atas adalah

$$CV=\frac{\left[\Gamma\left(1+\frac{2}{\gamma}\right)\Gamma\left(1-\frac{2}{\gamma}\right)-\left[\Gamma\left(1+\frac{1}{\gamma}\right)\Gamma\left(1-\frac{1}{\gamma}\right)\right]^{2}\right]^{1/2}}{\Gamma\left(1+\frac{1}{\gamma}\right)\Gamma\left(1-\frac{1}{\gamma}\right)},$$

yang hanya tergantung pada parameter γ.

Misalkan sekarang ada dua populasi berdistribusi log-logistik dengan parameter masing-masing adalah θ_1, γ_1 dan θ_2, γ_2. Misalkan juga bahwa median kedua populasi tersebut adalah M_1 dan M_2. Jelaslah bahwa kedua median tersebut dinyatakan sama ($M_1=M_2$) ketika hipotesis $H_0\colon\theta_1=\theta_2$ diterima. Interpretasi dari hipotesis ini tergantung pada apakah parameter γ_1 dan γ_2 sama atau tidak. Dengan demikian untuk uji hipotesis $H_0\colon\theta_1=\theta_2$ melawan $H_1\colon\theta_1\neq\theta_2$, ada dua kasus, yaitu kasus koefisien variasinya homogen ($\gamma_1=\gamma_2$), dan kasus koefisien variasinya heterogen ($\gamma_1\neq\gamma_2$).

Untuk kasus koefisien variasinya homogen, diasumsikan bahwa $\gamma_1=\gamma_2$. Dalam kasus ini, rata-rata dan varians dari distribusi log-logistik untuk kedua populasi mungkin saja berbeda, tetapi koefisien variasinya sama ($CV_1=CV_2$). Jika hipotesis $H_0\colon\theta_1=\theta_2$ diterima dalam kasus homogen, maka dapat disimpulkan bahwa kedua populasi identik. Untuk kasus koefisien variasinya heterogen, diasumsikan bahwa $\gamma_1\neq\gamma_2$. Jika hipotesis $H_0\colon\theta_1=\theta_2$ diterima dalam kasus heterogen, maka hanya dapat disimpulkan bahwa kedua median populasi identik.

Misalkan $x_{11},\cdots,x_{1n_1}$ dan $x_{21},\cdots,x_{2n_2}$ masing-masing menyatakan sampel-sampel saling bebas yang berukuran n_1 dan n_2 dari dua populasi log-logistik, $LL(\theta_1,\gamma_1)$ dan $LL(\theta_2,\gamma_2)$. Diasumsikan bahwa untuk setiap x_{ij} ada batas deteksi L_{ij}, untuk $i=1,2$ dan $j=1,2,\cdots,n_i$. Jika nilai pengamatannya terdeteksi, maka x_{ij} yang dicatat. Sedangkan jika nilai pengamatannya tidak terdeteksi ($<L_{ij}$), maka L_{ij} yang dicatat

(tersensor kiri). Misalkan untuk sampel i ada r_i pengamatan yang terdeteksi, sisanya $n_i - r_i$ pengamatan tidak terdeteksi.

Tahapan uji permutasi untuk hipotesis $H_0: \theta_1 = \theta_2$ melawan $H_1: \theta_1 \neq \theta_2$ adalah sebagai berikut (Mutaqin dan Kudus, 2014):

(1) Menghitung penaksir parameter θ_1 dan θ_2, yaitu $\hat{\theta}_1$ dan $\hat{\theta}_2$ berdasarkan data $(x_{ij}, L_{ij}; i = 1, 2$ dan $j = 1, 2, \cdots, n_i)$.

(2) Menghitung $\hat{\theta} = \hat{\theta}_1 - \hat{\theta}_2$.

(3) Mengambil sampel permutasi. Gabungkan data pengamatan yang terdeteksi dari sampel 1 dan 2, sehingga terbentuk data gabungan $X = (\boldsymbol{x}_1, \boldsymbol{x}_2) = (x_{11}, \cdots, x_{1r_1}, x_{21}, \cdots, x_{2r_2})$ untuk pengamatan yang terdeteksi berukuran $r = r_1 + r_2$. Gabungkan juga data pengamatan yang tidak terdeteksi dari sampel 1 dan 2, sehingga terbentuk data gabungan $L = (\boldsymbol{L}_1, \boldsymbol{L}_2) = (L_{1r_1+1}, \cdots, L_{1n_1}, L_{2r_2+1}, \cdots, L_{2n_2})$ untuk pengamatan yang tidak terdeteksi berukuran $s = n_1 + n_2 - (r_1 + r_2)$. Mengambil sampel acak berukuran r_1 tanpa pengembalian dari X, sebut saja sampel tersebut adalah $\boldsymbol{x}_1^*$. Sisanya yang tidak terambil dari X didefinisikan sebagai sampel $\boldsymbol{x}_2^*$. Kemudian mengambil sampel acak berukuran $n_1 - r_1$ tanpa pengembalian dari L, sebut saja sampel tersebut adalah $\boldsymbol{L}_1^*$. Sisanya yang tidak terambil dari L didefinisikan sebagai sampel $\boldsymbol{L}_2^*$.

(4) Menghitung penaksir parameter θ_1 dan θ_2, yaitu $\hat{\theta}_1^*$ dan $\hat{\theta}_2^*$ berdasarkan sampel permutasi $(\boldsymbol{x}_1^*, \boldsymbol{L}_1^*)$ dan $(\boldsymbol{x}_2^*, \boldsymbol{L}_2^*)$.

(5) Menghitung $\hat{\theta}^* = \hat{\theta}_1^* - \hat{\theta}_2^*$.

(6) Mengulang langkah (3) sampai (5) sebanyak M kali, sehingga diperoleh $\hat{\theta}^*$ sebanyak M buah.

(7) Menghitung ASL. Untuk uji dua pihak, yaitu proporsi banyaknya $|\hat{\theta}^*|$ yang nilainya lebih besar dari $|\hat{\theta}|$.

(8) Memutuskan apakah hipotesis nol diterima atau ditolak. Hipotesis nol diterima apabila nilai ASL lebih besar dari taraf arti (α) pengujian yang ditetapkan.

Hasil penelitian Mutaqin (2015) menunjukkan bahwa untuk data yang mengandung batas deteksi tunggal, uji permutasi akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi rendah dan perbedaan besar koefisien variasinya. Sedangkan untuk data yang mengandung batas deteksi ganda, uji permutasi akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi kedua sampel berbeda dan perbedaan besar koefisien variasi.

2.5. Pengujian Kesamaan Dua Median untuk Distribusi Log-logistik Menggunakan Informasi Fungsi Kemungkinan

2.5.1. Pengujian Kesamaan Dua Median

Misalkan ada dua populasi berdistribusi log-logistik dengan parameter masing-masing adalah θ_1, γ_1 dan θ_2, γ_2. Misalkan juga bahwa median kedua populasi tersebut adalah M_1 dan M_2. Jelaslah bahwa kedua median akan sama ($M_1 = M_2$) ketika hipotesis $H_0: \theta_1 = \theta_2$ diterima. Interpretasi dari hipotesis ini tergantung pada apakah parameter γ_1 dan γ_2 sama atau tidak.

Untuk kasus koefisien variasinya homogen, diasumsikan bahwa $\gamma_1 = \gamma_2$. Dalam kasus ini, rata-rata dan varians dari distribusi log-logistik untuk kedua populasi mungkin saja berbeda, tetapi koefisien variasinya sama ($CV_1 = CV_2$). Jika hipotesis $H_0: \theta_1 = \theta_2$ diterima dalam kasus homogen, maka dapat disimpulkan bahwa kedua populasi identik. Untuk kasus koefisien variasinya heterogen, diasumsikan bahwa $\gamma_1 \neq \gamma_2$. Jika hipotesis $H_0: \theta_1 = \theta_2$ diterima dalam kasus heterogen, maka hanya dapat disimpulkan bahwa kedua median populasi identik.

Dengan mengikuti hasil dari Stoline (1993), definisikan empat hipotesis berikut:

$$H_1: \theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma; \tag{2.11}$$

$$H_2: \theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma; \tag{2.12}$$

$$H_3: \theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2; \tag{2.13}$$

$$H_4: \theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2. \tag{2.14}$$

Empat pengujian hipotesis akan diusulkan, yaitu

uji hipotesis 1: H_1 lawan H_4; (2.15)

uji hipotesis 2: H_2 lawan H_4; (2.16)

uji hipotesis 3: H_1 lawan H_2; (2.17)

uji hipotesis 4: H_3 lawan H_4. (2.18)

Berikut ini akan direkomendasikan strategi pengujian melalui tiga tahapan, yaitu:

Tahap 1: menguji kehomogenan secara keseluruhan untuk dua populasi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi menggunakan uji hipotesis 1.

Tahap 2: menguji kehomogenan koefisien variasi menggunakan uji hipotesis 2.

Tahap 3: menguji kesamaan dua median menggunakan uji hipotesis 3 untuk kasus koefisien variasi homogen, atau menggunakan uji hipotesis 4 untuk kasus koefisien variasi tidak homogen.

Uji hipotesis 1 sampai 4 dilakukan dengan menggunakan informasi nilai fungsi kemungkinan dari hipotesis H_1, H_2, H_3 dan H_4. Berdasarkan nilai-nilai fungsi kemungkinan tersebut, dibentuk statistik uji chi-kuadrat asimtotik untuk semua uji hipotesis 1 sampai 4.

2.5.2. Pembentukan Fungsi Kemungkinan untuk Hipotesis

Berikut ini akan ditulis kembali beberapa notasi yang berkaitan dengan uji kesamaan dua median untuk distribusi log-logistik yang data pengamatannya mengandung pengamatan tidak terdeteksi. Misalkan $x_{11}, \cdots, x_{1n_1}$ dan $x_{21}, \cdots, x_{2n_2}$ masing-masing menyatakan sampel-sampel saling bebas yang berukuran n_1 dan n_2 dari dua populasi log-logistik, $LL(\theta_1, \gamma_1)$ dan $LL(\theta_2, \gamma_2)$. Diasumsikan bahwa untuk setiap x_{ij} ada batas deteksi L_{ij}, untuk $i = 1, 2$ dan $j = 1, 2, \cdots, n_i$. Jika nilai pengamatannya terdeteksi, maka x_{ij} yang dicatat. Sedangkan jika nilai pengamatannya tidak terdeteksi ($< L_{ij}$), maka L_{ij} yang dicatat (tersensor kiri). Misalkan untuk sampel i ada r_i pengamatan yang terdeteksi, sisanya $n_i - r_i$ pengamatan tidak terdeteksi.

Untuk menguji hipotesis 1 - 4 dalam Persamaan (2.15) - (2.18), dibutuhkan penaksir kemungkinan maksimum melalui algoritme EM dari parameter-parameter θ_1, $\gamma_1$dan θ_2, γ_2 dari empat kasus hipotesis H_1 - H_4 dalam Persamaan (2.11) - (2.14). Berikut ini akan diturunkan penaksir-penaksir tersebut untuk empat kasus hipotesis H_1 - H_4.

Kasus hipotesis H_1

Fungsi kemungkinan untuk kasus hipotesis H_1: $\theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma$ adalah

$$L_1(\theta, \gamma) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij}; \theta, \gamma) \prod_{j=r_i+1}^{n_i} F(L_{ij}; \theta, \gamma)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_1 = \ln L_1(\theta,\gamma) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta,\gamma) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta,\gamma)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma}\right]\right\}\right.$$

$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta - \ln\left[1+\left(L_{ij}/\theta\right)^{\gamma}\right]\right\}\right]$$

Kasus hipotesis H_2

Fungsi kemungkinan untuk kasus hipotesis H_2: $\theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma$ adalah

$$L_2(\theta_1,\theta_2,\gamma) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij};\theta_i,\gamma) \prod_{j=r_i+1}^{n_i} F(L_{ij};\theta_i,\gamma)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_2 = \ln L_2(\theta_1,\theta_2,\gamma) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta_i,\gamma) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta_i,\gamma)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta_i + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}\right]\right\}\right.$$

$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta_i - \ln\left[1+\left(L_{ij}/\theta_i\right)^{\gamma}\right]\right\}\right]$$

Kasus hipotesis H_3

Fungsi kemungkinan untuk kasus hipotesis H_3: $\theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2$ adalah

$$L_3(\theta,\gamma_1,\gamma_2) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij};\theta,\gamma_i) \prod_{j=r_i+1}^{n_i} F(L_{ij};\theta,\gamma_i)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_3 = \ln L_3(\theta,\gamma_1,\gamma_2) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta,\gamma_i) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta,\gamma_i)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i \ln\theta + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}\right]\right\}\right.$$

$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i \ln L_{ij} - \gamma_i \ln\theta - \ln\left[1 + \left(L_{ij}/\theta\right)^{\gamma_i}\right]\right\}\right]$$

Kasus hipotesis H_4

Fungsi kemungkinan untuk kasus hipotesis H_4: $\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$ adalah

$$L_4(\theta_1, \theta_2, \gamma_1, \gamma_2) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f\left(x_{ij}; \theta_i, \gamma_i\right) \prod_{j=r_i+1}^{n_i} F\left(L_{ij}; \theta_i, \gamma_i\right)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_4 = \ln L_4(\theta_1, \theta_2, \gamma_1, \gamma_2) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i} \ln f\left(x_{ij}; \theta_i, \gamma_i\right) + \sum_{j=r_i+1}^{n_i} \ln F\left(L_{ij}; \theta_i, \gamma_i\right)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i \ln\theta_i + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}\right]\right\}\right.$$

$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i \ln L_{ij} - \gamma_i \ln\theta_i - \ln\left[1 + \left(L_{ij}/\theta_i\right)^{\gamma_i}\right]\right\}\right]$$

2.5.3. Penaksiran Parameter

Berikut ini akan ditulis kembali beberapa notasi yang berkaitan dengan uji kesamaan dua median untuk distribusi log-logistik yang data pengamatannya mengandung pengamatan tidak terdeteksi. Misalkan $x_{11}, \cdots, x_{1n_1}$ dan $x_{21}, \cdots, x_{2n_2}$ masing-masing menyatakan sampel-sampel saling bebas yang berukuran n_1 dan n_2 dari dua populasi log-logistik, $LL(\theta_1, \gamma_1)$ dan $LL(\theta_2, \gamma_2)$. Diasumsikan bahwa untuk setiap x_{ij} ada batas deteksi L_{ij}, untuk $i = 1, 2$ dan $j = 1, 2, \cdots, n_i$. Jika nilai pengamatannya terdeteksi, maka x_{ij} yang dicatat. Sedangkan jika nilai pengamatannya tidak terdeteksi ($< L_{ij}$), maka L_{ij} yang dicatat (tersensor kiri). Misalkan untuk sampel i ada r_i pengamatan yang terdeteksi, sisanya $n_i - r_i$ pengamatan tidak terdeteksi.

Kasus hipotesis H_1

Fungsi kemungkinan untuk kasus hipotesis H_1: $\theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma$ adalah

$$L_1(\theta,\gamma) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij};\theta,\gamma)\prod_{j=r_i+1}^{n_i} F(L_{ij};\theta,\gamma)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_1 = \ln L_1(\theta,\gamma) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta,\gamma) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta,\gamma)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma}\right]\right\}\right.$$

$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta - \ln\left[1+\left(L_{ij}/\theta\right)^{\gamma}\right]\right\}\right]$$

Berdasarkan fungsi log-kemungkinan di atas, untuk menaksir kemungkinan maksimum melalui algoritme EM, perlu dihitung nilai ekspektasi dari $\ln(X_{ij})|X_{ij} < L_{ij}$ dan $\ln[1+(X_{ij}/\theta)^{\gamma}]|X_{ij} < L_{ij}$, untuk $i = 1,2$ dan $j = r_i+1, r_i+2, \cdots, n_i$. Dapat ditunjukkan bahwa nilai ekspektasi

$$E\left[\ln(X_{ij})|X_{ij} < L_{ij}\right]$$
$$= \ln\theta + \frac{\ln\left(F(L_{ij};\theta,\gamma)\right)}{\gamma}$$
$$+ \frac{\left[\frac{1}{F(L_{ij};\theta,\gamma)} - 1\right]\ln\left(1 - F(L_{ij};\theta,\gamma)\right)}{\gamma}, \tag{2.19}$$

dan

$$E\left[\ln[1+(X_{ij}/\theta)^{\gamma}]|X_{ij} < L_{ij}\right]$$
$$= \ln\left(1 + \frac{F(L_{ij};\theta,\gamma)}{1 - F(L_{ij};\theta,\gamma)}\right) + \frac{\ln\left(1 - F(L_{ij};\theta,\gamma)\right)}{F(L_{ij};\theta,\gamma)} + 1. \tag{2.20}$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln(x_{ij})$ dan $\ln[1+(x_{ij}/\theta)^{\gamma}]$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (2.19) dan (2.20). Misalkan $\theta^{(r)}$ dan $\gamma^{(r)}$ adalah taksiran parameter θ dan γ

pada iterasi ke-r, tahap-M adalah memaksimumkan fungsi kemungkinan berikut untuk memperoleh taksiran parameter θ dan γ pada iterasi ke-$r+1$

$$l_1' = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma E_{ij}^1 - \gamma\ln\theta - E_{ij}^2\right\}\right] \tag{2.21}$$

dimana E_{ij}^1 dan E_{ij}^2 masing-masing adalah ekspektasi yang ada pada Persamaan (2.19) dan (2.20), dengan θ dan γ diganti oleh $\theta^{(r)}$ dan $\gamma^{(r)}$. 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 (2.21) terhadap parameter θ dan γ masing-masing adalah

$$\frac{\partial l_1'}{\partial\theta} = -(n_1+n_2)\frac{\gamma}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma/\theta)(x_{ij}/\theta)^{\gamma}}{1+(x_{ij}/\theta)^{\gamma}}\right] \tag{2.22}$$

dan

$$\frac{\partial l_1'}{\partial\gamma} = \frac{(r_1+r_2)}{\gamma} - (n_1+n_2)\ln\theta + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln(x_{ij}/\theta)(x_{ij}/\theta)^{\gamma}}{1+(x_{ij}/\theta)^{\gamma}}\right\} + \sum_{j=r_i+1}^{n_i} E_{ij}^1\right]. \tag{2.23}$$

Kasus hipotesis H_2

Fungsi kemungkinan untuk kasus hipotesis H_2: $\theta_1 \neq \theta_2$, $\gamma_1 = \gamma_2 = \gamma$ adalah

$$L_2(\theta_1,\theta_2,\gamma) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij};\theta_i,\gamma)\prod_{j=r_i+1}^{n_i} F(L_{ij};\theta_i,\gamma)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_2 = \ln L_2(\theta_1,\theta_2,\gamma) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta_i,\gamma) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta_i,\gamma)\right]$$

$$= \sum_{i=1}^{2} \left[\sum_{j=1}^{r_i} \left\{ \ln\gamma - \gamma\ln\theta_i + (\gamma-1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}\right] \right\} \right.$$

$$\left. + \sum_{j=r_i+1}^{n_i} \left\{ \gamma\ln L_{ij} - \gamma\ln\theta_i - \ln\left[1 + \left(L_{ij}/\theta_i\right)^{\gamma}\right] \right\} \right]$$

Berdasarkan fungsi log-kemungkinan di atas, untuk menaksir kemungkinan maksimum melalui algoritme EM, perlu dihitung nilai ekspektasi dari $\ln(X_{ij})|X_{ij} < L_{ij}$ dan $\ln[1 + (X_{ij}/\theta_i)^{\gamma}]|X_{ij} < L_{ij}$, untuk $i = 1, 2$ dan $j = r_i + 1, r_i + 2, \cdots, n_i$. Dapat ditunjukkan bahwa nilai ekspektasi

$$E[\ln(X_{ij})|X_{ij} < L_{ij}]$$

$$= \ln\theta_i + \frac{\ln(F(L_{ij};\theta_i,\gamma))}{\gamma} + \frac{\left[\frac{1}{F(L_{ij};\theta_i,\gamma)} - 1\right]\ln(1 - F(L_{ij};\theta_i,\gamma))}{\gamma}, \quad (2.24)$$

dan

$$E[\ln[1 + (X_{ij}/\theta_i)^{\gamma}]|X_{ij} < L_{ij}]$$

$$= \ln\left(1 + \frac{F(L_{ij};\theta_i,\gamma)}{1 - F(L_{ij};\theta_i,\gamma)}\right) + \frac{\ln(1 - F(L_{ij};\theta_i,\gamma))}{F(L_{ij};\theta_i,\gamma)} + 1. \quad (2.25)$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln(x_{ij})$ dan $\ln[1 + (x_{ij}/\theta_i)^{\gamma}]$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (2.24) dan (2.25). Misalkan $\theta_1^{(r)}$, $\theta_2^{(r)}$ dan $\gamma^{(r)}$ adalah taksiran parameter θ_1, θ_2 dan γ pada iterasi ke-r, tahap-M adalah memaksimumkan fungsi kemungkinan berikut untuk memperoleh taksiran parameter θ_1, θ_2 dan γ pada iterasi ke-$r + 1$

$$l_2' = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta_i + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}\right]\right\}\right.$$
$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma E_{ij}^3 - \gamma\ln\theta_i - E_{ij}^4\right\}\right] \tag{2.26}$$

dimana E_{ij}^3 dan E_{ij}^4 masing-masing adalah ekspektasi yang ada pada Persamaan (2.24) dan (2.25), dengan θ_1, θ_2 dan γ diganti oleh $\theta_1^{(r)}$, $\theta_2^{(r)}$ dan $\gamma^{(r)}$. 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 (2.26) terhadap parameter θ dan γ masing-masing adalah

$$\frac{\partial l_2'}{\partial\theta_i} = -n_i\frac{\gamma}{\theta_i} + \sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}};\, i = 1,2 \tag{2.27}$$

dan

$$\frac{\partial l_2'}{\partial\gamma} = \frac{(r_1+r_2)}{\gamma} - (n_1\ln\theta_1 + n_2\ln\theta_2)$$
$$+\sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln(x_{ij}/\theta_i)\,(x_{ij}/\theta_i)^{\gamma}}{1+(x_{ij}/\theta_i)^{\gamma}}\right\} + \sum_{j=r_i+1}^{n_i} E_{ij}^3\right]. \tag{2.28}$$

Kasus hipotesis H_3

Fungsi kemungkinan untuk kasus hipotesis H_3: $\theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2$ adalah

$$L_3(\theta,\gamma_1,\gamma_2) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij};\theta,\gamma_i)\prod_{j=r_i+1}^{n_i} F(L_{ij};\theta,\gamma_i)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_3 = \ln L_3(\theta,\gamma_1,\gamma_2) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij};\theta,\gamma_i) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij};\theta,\gamma_i)\right]$$

$$= \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i\ln\theta + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}\right]\right\}\right.$$
$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i\ln L_{ij} - \gamma_i\ln\theta - \ln\left[1 + \left(L_{ij}/\theta\right)^{\gamma_i}\right]\right\}\right]$$

Berdasarkan fungsi log-kemungkinan di atas, untuk menaksir kemungkinan maksimum melalui algoritme EM, perlu dihitung nilai ekspektasi dari $\ln(X_{ij})|X_{ij} < L_{ij}$ dan $\ln\left[1 + \left(X_{ij}/\theta\right)^{\gamma_i}\right]|X_{ij} < L_{ij}$, untuk $i = 1, 2$ dan $j = r_i + 1, r_i + 2, \cdots, n_i$. Dapat ditunjukkan bahwa nilai ekspektasi

$$\begin{aligned} &E\left[\ln(X_{ij})|X_{ij} < L_{ij}\right] \\ &\qquad = \ln\theta + \frac{\ln\left(F(L_{ij};\theta,\gamma_i)\right)}{\gamma_i} \\ &\qquad + \frac{\left[\frac{1}{F(L_{ij};\theta,\gamma_i)} - 1\right]\ln\left(1 - F(L_{ij};\theta,\gamma_i)\right)}{\gamma_i}, \end{aligned} \tag{2.29}$$

dan

$$\begin{aligned} &E\left[\ln\left[1 + \left(X_{ij}/\theta\right)^{\gamma_i}\right]|X_{ij} < L_{ij}\right] \\ &\qquad = \ln\left(1 + \frac{F(L_{ij};\theta,\gamma_i)}{1 - F(L_{ij};\theta,\gamma_i)}\right) + \frac{\ln\left(1 - F(L_{ij};\theta,\gamma_i)\right)}{F(L_{ij};\theta,\gamma_i)} + 1. \end{aligned} \tag{2.30}$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln(x_{ij})$ dan $\ln\left[1 + \left(x_{ij}/\theta\right)^{\gamma_i}\right]$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (2.29) dan (2.30). Misalkan $\theta^{(r)}$, $\gamma_1^{(r)}$ dan $\gamma_2^{(r)}$ adalah taksiran parameter θ, γ_1, dan γ_2 pada iterasi ke-r, tahap-M adalah memaksimumkan fungsi kemungkinan berikut untuk memperoleh taksiran parameter θ, γ_1, dan γ_2 pada iterasi ke-$r + 1$

$$l_3' = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i\ln\theta + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}\right]\right\}\right.$$
$$\left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i E_{ij}^5 - \gamma_i\ln\theta - E_{ij}^6\right\}\right] \quad (2.31)$$

dimana E_{ij}^5 dan E_{ij}^6 masing-masing adalah ekspektasi yang ada pada Persamaan (2.29) dan (2.30), dengan θ, γ_1, dan γ_2 diganti oleh $\theta^{(r)}$, $\gamma_1^{(r)}$ dan $\gamma_2^{(r)}$. 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 (2.31) terhadap parameter θ, γ_1, dan γ_2 masing-masing adalah

$$\frac{\partial l_3'}{\partial\theta} = -\frac{(n_1\gamma_1 + n_2\gamma_2)}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma_i/\theta)(x_{ij}/\theta)^{\gamma_i}}{1 + (x_{ij}/\theta)^{\gamma_i}}\right] \quad (2.32)$$

dan

$$\frac{\partial l_3'}{\partial\gamma_i} = \frac{r_i}{\gamma_i} - n_i\ln\theta + \sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}{1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}\right\}$$
$$+ \sum_{j=r_i+1}^{n_i} E_{ij}^5 ; i = 1, 2 \quad (2.33)$$

Kasus hipotesis H_4

Fungsi kemungkinan untuk kasus hipotesis H_4: $\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$ adalah

$$L_4(\theta_1, \theta_2, \gamma_1, \gamma_2) = \prod_{i=1}^{2}\left[\prod_{j=1}^{r_i} f(x_{ij}; \theta_i, \gamma_i) \prod_{j=r_i+1}^{n_i} F(L_{ij}; \theta_i, \gamma_i)\right]$$

Fungsi log-kemungkinan-nya adalah

$$l_4 = \ln L_4(\theta_1, \theta_2, \gamma_1, \gamma_2) = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\ln f(x_{ij}; \theta_i, \gamma_i) + \sum_{j=r_i+1}^{n_i}\ln F(L_{ij}; \theta_i, \gamma_i)\right]$$

$$= \sum_{i=1}^{2} \left[\sum_{j=1}^{r_i} \left\{ \ln \gamma_i - \gamma_i \ln \theta_i + (\gamma_i - 1) \ln x_{ij} - 2 \ln \left[1 + \left(\frac{x_{ij}}{\theta_i} \right)^{\gamma_i} \right] \right\} \right.$$

$$\left. + \sum_{j=r_i+1}^{n_i} \left\{ \gamma_i \ln L_{ij} - \gamma_i \ln \theta_i - \ln\left[1 + \left(L_{ij}/\theta_i\right)^{\gamma_i}\right] \right\} \right]$$

Berdasarkan fungsi log-kemungkinan di atas, untuk menaksir kemungkinan maksimum melalui algoritme EM, perlu dihitung nilai ekspektasi dari $\ln(X_{ij})|X_{ij} < L_{ij}$ dan $\ln[1 + (X_{ij}/\theta_i)^{\gamma_i}]|X_{ij} < L_{ij}$, untuk $i = 1, 2$ dan $j = r_i + 1, r_i + 2, \cdots, n_i$. Dapat ditunjukkan bahwa nilai ekspektasi

$$E[\ln(X_{ij})|X_{ij} < L_{ij}]$$
$$= \ln \theta_i + \frac{\ln\left(F(L_{ij}; \theta_i, \gamma_i)\right)}{\gamma_i} \qquad (2.34)$$
$$+ \frac{\left[\frac{1}{F(L_{ij}; \theta_i, \gamma_i)} - 1\right] \ln\left(1 - F(L_{ij}; \theta_i, \gamma_i)\right)}{\gamma_i},$$

dan

$$E[\ln[1 + (X_{ij}/\theta_i)^{\gamma_i}]|X_{ij} < L_{ij}]$$
$$= \ln\left(1 + \frac{F(L_{ij}; \theta_i, \gamma_i)}{1 - F(L_{ij}; \theta_i, \gamma_i)}\right) + \frac{\ln\left(1 - F(L_{ij}; \theta_i, \gamma_i)\right)}{F(L_{ij}; \theta_i, \gamma_i)} + 1. \qquad (2.35)$$

Dengan demikian tahap-E dalam algoritme EM adalah mengganti $\ln(x_{ij})$ dan $\ln[1 + (x_{ij}/\theta_i)^{\gamma_i}]$ dalam fungsi log-kemungkinan untuk data lengkap masing-masing oleh Persamaan (2.34) dan (2.35). Misalkan $\theta_1^{(r)}$, $\theta_2^{(r)}$, $\gamma_1^{(r)}$ dan $\gamma_2^{(r)}$ adalah taksiran parameter θ_1, θ_2, γ_1, dan γ_2 pada iterasi ke-r, tahap-M adalah memaksimumkan fungsi kemungkinan berikut untuk memperoleh taksiran parameter θ_1, θ_2, γ_1, dan γ_2 pada iterasi ke-$r + 1$

$$l_4' = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i\ln\theta_i + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i E_{ij}^7 - \gamma_i\ln\theta_i - E_{ij}^8\right\}\right] \tag{2.36}$$

dimana E_{ij}^7 dan E_{ij}^8 masing-masing adalah ekspektasi yang ada pada Persamaan (2.34) dan (2.35), dengan θ_1, θ_2, γ_1, dan γ_2 diganti oleh $\theta_1^{(r)}$, $\theta_2^{(r)}$, $\gamma_1^{(r)}$ dan $\gamma_2^{(r)}$. 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 (2.36) terhadap parameter θ_1, θ_2, γ_1, dan γ_2 masing-masing adalah

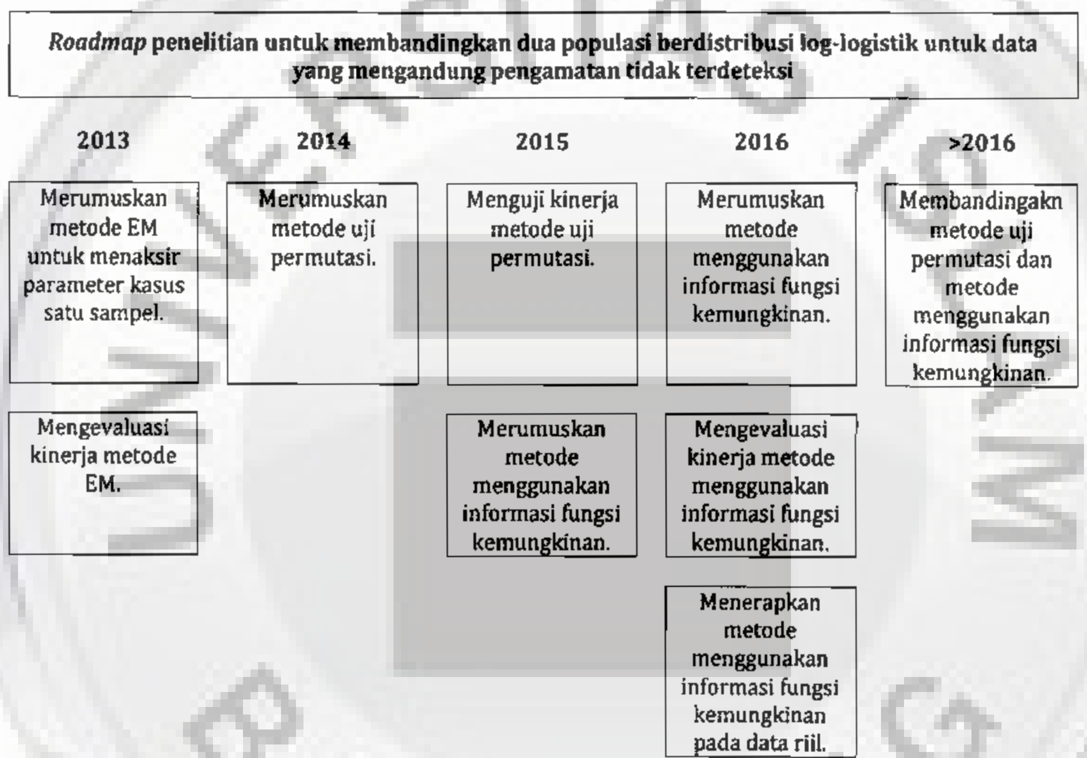
$$\frac{\partial l_4'}{\partial\theta_i} = -n_i\frac{\gamma_i}{\theta_i} + \sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}};\, i = 1, 2 \tag{2.37}$$

dan

$$\frac{\partial l_4'}{\partial\gamma_i} = \frac{r_i}{\gamma_i} - n_i\ln\theta_i + \sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}\right\} + \sum_{j=r_i+1}^{n_i} E_{ij}^7;\, i = 1, 2. \tag{2.38}$$

2.6. *Roadmap*

Mutaqin dkk. (2013) telah melakukan penelitian penaksiran parameter populasi satu sampel yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi menggunakan algoritme EM. Mutaqin dan Kudus (2014) telah membahas uji permutasi untuk membandingkan dua populasi berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi dengan jalan membandingkan kedua median populasinya. Hasil kinerja dari uji permutasi di atas telah dikaji oleh Mutaqin (2015). Di tahun 2015, telah dikaji metode lain untuk membandingkan dua populasi berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi menggunakan informasi fungsi kemungkinan. Namun demikian, kajiannya belum tuntas dan rencananya akan diselesaikan di tahun 2016. Di tahun 2016, metode yang menggunakan informasi fungsi kemungkinan akan dievaluasi kinerjanya menggunakan simulasi Monte Carlo dan akan diaplikasikan pada data riil. Setelah tahun 2016, metode yang menggunakan uji permutasi dan yang menggunakan informasi fungsi kemungkinan akan dibandingkan. *Roadmap* penelitian secara ringkas diilustrasikan dalam Gambar 2.1.



Gambar 2.1. *Roadmap* Penelitian

BAB 3. TUJUAN DAN MANFAAT PENELITIAN

3.1. Tujuan Penelitian

Tujuan dari penelitian ini:

a. Untuk merumuskan metode berdasarkan fungsi kemungkinan untuk membandingkan dua sampel berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi.

b. Untuk mengevaluasi kinerja dari metode berdasarkan fungsi kemungkinan untuk membandingkan dua sampel berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi berdasarkan kriteria kuasa uji dan kesalahan tipe I menggunakan simulasi.

c. Untuk menerapkan metode berdasarkan fungsi kemungkinan untuk membandingkan dua sampel berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi pada data lingkungan.

3.2. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat bermanfaat dalam:

a. Menghasilkan suatu alternatif prosedur statistika untuk menangani kasus data sampel yang mengandung pengamatan tidak terdeteksi.

b. Hasil penelitian ini diharapkan menjadi rujukan bagi peneliti dan praktisi dalam penelitiannya yang berkaitan dengan kasus data sampel yang mengandung pengamatan tidak terdeteksi.

BAB 4. METODE PENELITIAN

4.1. Metode

Dalam penelitian ini akan dirumuskan metode perbandingan dua sampel saling bebas berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi menggunakan informasi fungsi kemungkinan. Tahapan penelitian yang akan dilakukan adalah:

a. Membangun uji hipotesis untuk membandingkan dua populasi yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi.
b. Merumuskan statistik uji untuk uji hipotesis di atas dengan jalan membentuk semua fungsi kemungkinan dalam hipotesis-hipotesisnya. Dalam tahap ini dilibatkan pendugaan data tidak terdeteksi dengan menggunakan metode EM yang sudah dibuat oleh Mutaqin dkk. (2013).

Metode yang dirumuskan akan dievaluasi kinerjanya berdasarkan kriteria kuasa uji dan kesalahan tipe I menggunakan simulasi Monte Carlo menggunakan perangkat lunak statistika (misalnya R, Matlab atau SAS).

4.2. Data

Metode yang diusulkan akan coba diterapkan penggunaanya pada data kadar carbon monoksida (CO) hasil uji emisi gas buang kendaraan bermotor dua pabrikan kendaraan *A* dan *B* di Kota Bandung. Tabel 4.1 menyajikan data kadar CO dalam persen untuk kendaraan-kendaraan kedua pabrikan tersebut.

4.3. Luaran

Luaran yang diharapkan dari penelitian ini adalah:

(1) Terbangunnya metode perbandingan dua populasi yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi dengan menggunakan informasi fungsi kemungkinan.
(2) Dihasilkannya kinerja metode perbandingan dua populasi yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi dengan menggunakan informasi fungsi kemungkinan.

Luaran-luaran di atas akan ditulis dalam satu atau dua artikel ilmiah yang akan dipublikasikan dalam prosiding seminar internasional dan atau jurnal ilmiah bereputasi internasional.

Tabel 3.1. Data Kadar CO (dalam %)

Pabrikan *A*	Pabrikan *B*
0,00*	0,00*
0,01	0,00*
0,01	0,01
0,01	0,01
0,02	0,01
0,02	0,01
0,03	0,01
0,05	0,02
0,09	0,02
2,21	0,02
2,29	0,17
	0,18
	0,47
	1,50
	2,53

*tidak terdeteksi (batas deteksi 0,01)

BAB 5. HASIL DAN LUARAN YANG DICAPAI

5.1. Penaksiran Parameter

Misalkan $x_{11}, \cdots, x_{1n_1}$ dan $x_{21}, \cdots, x_{2n_2}$ masing-masing menyatakan sampel-sampel saling bebas yang berukuran n_1 dan n_2 dari dua populasi log-logistik, $LL(\theta_1, \gamma_1)$ dan $LL(\theta_2, \gamma_2)$. Diasumsikan bahwa untuk setiap x_{ij} ada batas deteksi L_{ij}, untuk $i = 1, 2$ dan $j = 1, 2, \cdots, n_i$. Jika nilai pengamatannya terdeteksi, maka x_{ij} yang dicatat. Sedangkan jika nilai pengamatannya tidak terdeteksi ($< L_{ij}$), maka L_{ij} yang dicatat (tersensor kiri). Misalkan untuk sampel i ada r_i pengamatan yang terdeteksi, sisanya $n_i - r_i$ pengamatan tidak terdeteksi. Dalam menguji kesamaan dua median untuk distribusi log-logistik yang data pengamatannya mengandung pengamatan tidak terdeteksi, diperlukan nilai taksiran parameter untuk setiap kasus hipotesis

H_1: $\theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma$;

H_2: $\theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma$;

H_3: $\theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2$;

H_4: $\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$.

Untuk hipotesis H_1 ada dua parameter yang perlu ditaksir, yaitu θ dan γ. Kedua parameter tersebut ditaksir menggunakan algoritme EM melalui metode Newton-Raphson. Nilai taksiran parameter θ dan γ diperoleh sebagai solusi dari dua persamaan berikut

$$-(n_1 + n_2)\frac{\gamma}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i} \frac{2\left(\frac{\gamma}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma}}{1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma}}\right] = 0,$$

dan

$$\frac{(r_1 + r_2)}{\gamma} - (n_1 + n_2)\ln\theta$$
$$+ \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma}}{1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma}}\right\} + \sum_{j=r_i+1}^{n_i} E_{ij}^{1}\right] = 0.$$

Untuk hipotesis H_2 ada tiga parameter yang perlu ditaksir, yaitu θ_1, θ_2 dan γ. Ketiga parameter tersebut ditaksir menggunakan algoritme EM melalui metode Newton-Raphson. Nilai taksiran parameter θ_1, θ_2 dan γ diperoleh sebagai solusi dari tiga persamaan berikut

$$-n_i\frac{\gamma}{\theta_i}+\sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}=0; i=1,2$$

dan

$$\frac{(r_1+r_2)}{\gamma}-(n_1\ln\theta_1+n_2\ln\theta_2)$$
$$+\sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij}-\frac{2\ln\left(\frac{x_{ij}}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}\right\}+\sum_{j=r_i+1}^{n_i}E_{ij}^{3}\right]=0.$$

Untuk hipotesis H_3 ada tiga parameter yang perlu ditaksir, yaitu θ, γ_1, dan γ_2. Ketiga parameter tersebut ditaksir menggunakan algoritme EM melalui metode Newton-Raphson. Nilai taksiran parameter θ, γ_1, dan γ_2 diperoleh sebagai solusi dari tiga persamaan berikut

$$-\frac{(n_1\gamma_1+n_2\gamma_2)}{\theta}+\sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma_i}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}\right]=0,$$

dan

$$\frac{r_i}{\gamma_i}-n_i\ln\theta+\sum_{j=1}^{r_i}\left\{\ln x_{ij}-\frac{2\ln\left(\frac{x_{ij}}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}\right\}+\sum_{j=r_i+1}^{n_i}E_{ij}^{5}=0; i=1,2.$$

Untuk hipotesis H_4 ada empat parameter yang perlu ditaksir, yaitu θ_1, θ_2, γ_1, dan γ_2. Keempat parameter tersebut ditaksir menggunakan algoritme EM melalui metode Newton-Raphson. Nilai taksiran parameter θ_1, θ_2, γ_1, dan γ_2 diperoleh sebagai solusi dari empat persamaan berikut

$$-n_i \frac{\gamma_i}{\theta_i} + \sum_{j=1}^{r_i} \frac{2\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}} = 0; i = 1, 2$$

dan

$$\frac{r_i}{\gamma_i} - n_i \ln \theta_i + \sum_{j=1}^{r_i} \left\{ \ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}} \right\} + \sum_{j=r_i+1}^{n_i} E_{ij}^{7} = 0; i = 1, 2.$$

Program Matlab untuk menaksir parameter-parameter di atas disajikan dalam Lampiran.

5.2. Prosedur Pengujian yang Diusulkan

Dengan mengikuti hasil dari Stoline (1993), definisikan empat hipotesis berikut:

$H_1: \theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma;$

$H_2: \theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma;$

$H_3: \theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2;$

$H_4: \theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2.$

Empat pengujian hipotesis akan diusulkan, yaitu

uji hipotesis 1: H_1 lawan H_4; (4.1)

uji hipotesis 2: H_2 lawan H_4; (4.2)

uji hipotesis 3: H_1 lawan H_2; (4.3)

uji hipotesis 4: H_3 lawan H_4. (4.4)

Berikut ini akan direkomendasikan strategi pengujian melalui tiga tahapan, yaitu:

Tahap 1: menguji kehomogenan secara keseluruhan untuk dua populasi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi menggunakan uji hipotesis 1.

Tahap 2: menguji kehomogenan koefisien variasi menggunakan uji hipotesis 2.

Tahap 3: menguji kesamaan dua median menggunakan uji hipotesis 3 untuk kasus koefisien variasi homogen, atau menggunakan uji hipotesis 4 untuk kasus koefisien variasi tidak homogen.

Uji hipotesis 1 sampai 4 dilakukan dengan menggunakan informasi nilai fungsi kemungkinan dari hipotesis H_1, H_2, H_3 dan H_4. Berdasarkan nilai-nilai fungsi kemungkinan tersebut, dibentuk statistik uji chi-kuadrat asimtotik untuk semua uji hipotesis 1 sampai 4.

Fungsi log-kemungkinan untuk kasus hipotesis $H_1: \theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma$ adalah

$$l_1 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta - \ln\left[1+\left(L_{ij}/\theta\right)^{\gamma}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_1 terhadap parameter θ dan γ masing-masing adalah

$$\frac{\partial l_1}{\partial\theta} = -(n_1+n_2)\frac{\gamma}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma/\theta)\left(x_{ij}/\theta\right)^{\gamma}}{1+\left(x_{ij}/\theta\right)^{\gamma}} + \sum_{j=r_i+1}^{n_i}\frac{(\gamma/\theta)\left(L_{ij}/\theta\right)^{\gamma}}{1+\left(L_{ij}/\theta\right)^{\gamma}}\right] \quad (4.5)$$

dan

$$\frac{\partial l_1}{\partial\gamma} = \frac{(r_1+r_2)}{\gamma} - (n_1+n_2)\ln\theta + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(x_{ij}/\theta\right)\left(x_{ij}/\theta\right)^{\gamma}}{1+\left(x_{ij}/\theta\right)^{\gamma}}\right\} + \sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(L_{ij}/\theta\right)\left(L_{ij}/\theta\right)^{\gamma}}{1+\left(L_{ij}/\theta\right)^{\gamma}}\right\}\right]. \quad (4.6)$$

Fungsi log-kemungkinan untuk kasus hipotesis $H_2: \theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma$ adalah

$$l_2 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta_i + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta_i - \ln\left[1+\left(L_{ij}/\theta_i\right)^{\gamma}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_2 terhadap parameter θ_1, θ_2 dan γ masing-masing adalah

$$\frac{\partial l_2}{\partial \theta_i} = -n_i \frac{\gamma}{\theta_i} + \sum_{j=1}^{r_i} \frac{2\left(\frac{\gamma}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}{1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}} + \sum_{j=r_i+1}^{n_i} \frac{(\gamma/\theta_i)\left(L_{ij}/\theta_i\right)^{\gamma}}{1 + \left(L_{ij}/\theta_i\right)^{\gamma}}; i = 1, 2 \tag{4.7}$$

dan

$$\begin{aligned}\frac{\partial l_2}{\partial \gamma} = \frac{(r_1 + r_2)}{\gamma} - (n_1 \ln \theta_1 + n_2 \ln \theta_2) \\ + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(x_{ij}/\theta_i\right)\left(x_{ij}/\theta_i\right)^{\gamma}}{1 + \left(x_{ij}/\theta_i\right)^{\gamma}}\right\}\right. \\ \left. + \sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(L_{ij}/\theta_i\right)\left(L_{ij}/\theta_i\right)^{\gamma}}{1 + \left(L_{ij}/\theta_i\right)^{\gamma}}\right\}\right].\end{aligned} \tag{4.8}$$

Fungsi log-kemungkinan untuk kasus hipotesis H_3: $\theta_1 = \theta_2 = \theta$, $\gamma_1 \neq \gamma_2$ adalah

$$\begin{aligned}l_3 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln \gamma_i - \gamma_i \ln \theta + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}\right]\right\}\right. \\ \left. + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i \ln L_{ij} - \gamma_i \ln \theta - \ln\left[1 + \left(L_{ij}/\theta\right)^{\gamma_i}\right]\right\}\right]\end{aligned}$$

Turunan pertama dari fungsi log-kemungkinan l_3 terhadap parameter θ, γ_1, dan γ_2 masing-masing adalah

$$\begin{aligned}\frac{\partial l_3}{\partial \theta} = -\frac{(n_1\gamma_1 + n_2\gamma_2)}{\theta} \\ + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma_i/\theta)\left(x_{ij}/\theta\right)^{\gamma_i}}{1 + \left(x_{ij}/\theta\right)^{\gamma_i}} + \sum_{j=r_i+1}^{n_i}\frac{(\gamma_i/\theta)\left(L_{ij}/\theta\right)^{\gamma_i}}{1 + \left(L_{ij}/\theta\right)^{\gamma_i}}\right]\end{aligned} \tag{4.9}$$

dan

$$\frac{\partial l_3}{\partial \gamma_i} = \frac{r_i}{\gamma_i} - n_i \ln \theta + \sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}{1 + \left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}\right\} + \tag{4.10}$$

$$+\sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij}-\frac{\ln(L_{ij}/\theta)\left(L_{ij}/\theta\right)^{\gamma_i}}{1+\left(L_{ij}/\theta\right)^{\gamma_i}}\right\};i=1,2$$

Fungsi log-kemungkinan untuk kasus hipotesis H_4: $\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$ adalah

$$l_4=\sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i-\gamma_i\ln\theta_i+(\gamma_i-1)\ln x_{ij}-2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}\right]\right\}\right.$$
$$\left.+\sum_{j=r_i+1}^{n_i}\left\{\gamma_i\ln L_{ij}-\gamma_i\ln\theta_i-\ln\left[1+\left(L_{ij}/\theta_i\right)^{\gamma_i}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_4 terhadap parameter θ_1, θ_2, γ_1, dan γ_2 masing-masing adalah

$$\frac{\partial l_4}{\partial\theta_i}=-n_i\frac{\gamma_i}{\theta_i}+\sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}+\sum_{j=r_i+1}^{n_i}\frac{\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}};i=1,2 \tag{4.11}$$

dan

$$\frac{\partial l_4}{\partial\gamma_i}=\frac{r_i}{\gamma_i}-n_i\ln\theta_i+\sum_{j=1}^{r_i}\left\{\ln x_{ij}-\frac{2\ln\left(\frac{x_{ij}}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}\right\}$$
$$+\sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij}-\frac{\ln\left(\frac{L_{ij}}{\theta_i}\right)\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}{1+\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}\right\};i=1,2. \tag{4.12}$$

Uji chi-square asimtotik dapat digunakan untuk menguji hipotesis 1 sampai uji hipotesis 4 dengan statistika ujinya masing-masing adalah

$$\chi_1^2=-2\left(\hat{l}_1-\hat{l}_4\right)\sim\chi^2_{2,1-\alpha},$$
$$\chi_2^2=-2\left(\hat{l}_2-\hat{l}_4\right)\sim\chi^2_{1,1-\alpha},$$
$$\chi_3^2=-2\left(\hat{l}_1-\hat{l}_2\right)\sim\chi^2_{1,1-\alpha},$$
$$\chi_4^2=-2\left(\hat{l}_3-\hat{l}_4\right)\sim\chi^2_{1,1-\alpha},$$

dimana $\chi^2_{r,1-\alpha}$ adalah kuantil ke-$100(1-\alpha)$ dari distribusi chi-square dengan derajat bebas r. Sementara itu $\hat{l}_1$, $\hat{l}_2$, $\hat{l}_3$, dan $\hat{l}_4$ masing-masing merupakan taksiran kemungkinan maksimum untuk nilai fungsi log-kemungkinan pada hipotesis H_1, H_2, H_3, dan H_4.

5.3. Kinerja Metode Usulan

Kinerja dari metode yang diusulkan akan dievaluasi dengan menggunakan simulasi Monte Carlo. Evaluasi kinerjanya didasarkan pada kesalahan tipe I empirik dan kuasa uji empirik dari hasil simulasi Monte Carlo. Lampiran 2 menyajikan program simulasi Monte Carlo. Beberapa kasus akan dicobakan dalam simulasi untuk melihat pengaruh ukuran sampel, persentase pengamatan yang tidak terdeteksi, dan kombinasi nilai parameter dari distribusi untuk setiap sampel. Kasus-kasus data tersebut disajikan dalam Tabel 5.1.

Tabel 5.1. Beberapa kasus data dalam simulasi Monte Carlo

Kasus	Ukuran Sampel	Persentase Pengamatan Tidak Terdeteksi	Kombinasi Parameter
1	20	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
2	50	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
3	100	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
4	20	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
5	50	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
6	100	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -6$
7	20	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
8	50	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
9	100	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
10	20	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
11	50	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
12	100	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -6$
13	20	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
14	50	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
15	100	0,1	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
16	20	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
17	50	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
18	100	0,3	$\alpha_1 = 4; \beta_1 = -6$ dan $\alpha_2 = 4; \beta_2 = -4$
19	20	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$
20	50	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$
21	100	0,1	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$
22	20	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$
23	50	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$
24	100	0,3	$\alpha_1 = 2; \beta_1 = -3$ dan $\alpha_2 = 4; \beta_2 = -4$

Dalam Tabel 5.1, kasus 1 sampai dengan kasus 6 adalah kasus kesamaan dua median populasi dan kesamaan dua koefisien variasi populasi. Kasus 7 sampai dengan kasus 12 adalah kasus kesamaan dua median populasi dan perbedaan dua koefisien variasi populasi. Kasus 13 sampai dengan kasus 18 adalah kasus perbedaan dua median populasi dan kesamaan dua koefisien variasi populasi. Sedangkan kasus 19 sampai dengan kasus 24 adalah kasus perbedaan dua median populasi dan perbedaan dua koefisien variasi populasi. Kasus 1 sampai dengan kasus 12 dipakai untuk melihat kesalahan tipe I empirik dari pengujian yang diusulkan. Sedangkan kasus 13 sampai dengan kasus 24 dipakai untuk melihat kuasa uji empiric dari pengujian yang diusulkan.

Kesalahan tipe I empirik dan kuasa uji empirik untuk uji permutasi kesamaan dua median dari dua populasi berdistribusi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi untuk kasus data yang mengandung batas deteksi tunggal masing-masing disajikan dalam Tabel 5.2 dan 5.3.

Tabel 5.2. Kesalahan tipe I empirik
untuk kasus data yang dicobakan

Kasus	Ukuran Sampel	Persentase Pengamatan Tidak Terdeteksi	Kesalahan Tipe I Empirik
1	20	0,1	0,062
2	50	0,1	0,049
3	100	0,1	0,047
4	20	0,3	0,062
5	50	0,3	0,067
6	100	0,3	0,047
7	20	0,1	0,007
8	50	0,1	0,007
9	100	0,1	0,027
10	20	0,3	0,100
11	50	0,3	0,020
12	100	0,3	0,060

Hasil simulasi Monte Carlo dalam Tabel 5.2 dan 5.3 menunjukkan bahwa metode yang diusulkan memiliki kinerja yang bagus terutama untuk ukuran sampel yang besar. Hal ini terlihat dari dengan semakin besarnya ukuran sampel maka nilai kesalahan tipe I empiriknya dekat ke nilai taraf nyata yang ditetapkan, yaitu 5% dan nilai kuasa uji empiriknya semakin menuju 1.

Tabel 5.3. Kuasa uji empirik
untuk kasus data yang dicobakan

Kasus	Ukuran Sampel	Persentase Pengamatan Tidak Terdeteksi	Kuasa uji Empirik
1	20	0,1	0,88
2	50	0,1	1,00
3	100	0,1	1,00
4	20	0,3	0,66
5	50	0,3	1,00
6	100	0,3	1,00
7	20	0,1	0,91
8	50	0,1	1,00
9	100	0,1	1,00
10	20	0,3	0,28
11	50	0,3	0,80
12	100	0,3	1,00

5.4. Contoh Numerik

Dalam bagian ini akan diberikan contoh kasus penerapan metode yang diusulkan pada data kadar carbon monoksida (CO) hasil uji emisi gas buang kendaraan bermotor dua pabrikan kendaraan *A* dan *B* di Kota Bandung. Tabel 5.1 menyajikan data kadar CO dalam persen untuk kendaraan-kendaraan kedua pabrikan tersebut.

Tabel 1. Data Kadar CO (dalam %)

Pabrikan *A*	Pabrikan *B*
0,00*	0,00*
0,01	0,00*
0,01	0,01
0,01	0,01
0,02	0,01
0,02	0,01
0,03	0,01
0,05	0,02
0,09	0,02
2,21	0,02
2,29	0,17
	0,18
	0,47
	1,50
	2,53

*tidak terdeteksi (batas deteksi 0,01)

Dengan menerapkan metode yang diusulkan diperoleh nilai p-value untuk uji hipotesis 1, uji hipotesis 2 dan uji hipotesis 3 masing-masing sebesar 0,5925; 0,6162; dan 0,9615. Berdasarkan nilai-nilai p-value tersebut dapat disimpulkan bahwa secara statistik median dan distribusi kadar CO untuk kendaraan-kendaraan pabrikan *A* dan *B* di Kota Bandung adalah sama.

5.5. Luaran yang Dicapai

Luaran yang telah dicapai dari hasil penelitian ini adalah:

(1) Terbangunnya metode perbandingan dua populasi yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi dengan menggunakan informasi fungsi kemungkinan.

(2) Dihasilkannya kinerja metode perbandingan dua populasi yang berdistribusi log-logistik yang mengandung pengamatan tidak terdeteksi dengan menggunakan informasi fungsi kemungkinan.

Luaran-luaran di atas telah ditulis dalam satu artikel ilmiah yang akan dipublikasikan dalam jurnal ilmiah bereputasi internasional. Draft artikel ilmiahnya disajikan dalam Lampiran.

BAB 6. KESIMPULAN DAN SARAN

6.1. Kesimpulan

Kesimpulan yang dapat ditarik dari hasil penelitian ini adalah:

a) Telah dirumuskan metode uji parametrik menggunakan informasi fungsi kemungkinan untuk membandingkan dua populasi berdistribusi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi. Uji parametrik ini dapat menjawab pertanyaan apakah hasil pengujiannya termasuk untuk kasus koefisien variasi sama atau berbeda. Dalam penelitian ini yang diuji adalah kesamaan median kedua populasi tersebut. Program MATLAB telah dibuat untuk pengujian di atas.

b) Hasil simulasi Monte Carlo menunjukkan bahwa metode yang diusulkan memiliki kinerja yang bagus terutama untuk ukuran sampel yang besar.

c) Hasil penerapan metode yang diusulkan pada data kadar carbon monoksida (CO) hasil uji emisi gas buang kendaraan bermotor dua pabrikan kendaraan *A* dan *B* di Kota Bandung menunjukkan bahwa secara statistik median dan distribusi kadar CO untuk kendaraan-kendaraan pabrikan *A* dan *B* di Kota Bandung adalah sama.

6.2. Saran

Saran yang dapat diberikan sehubungan dengan hasil penelitian ini adalah:

a) Disarankan untuk peneliti dan praktisi yang berhadapan dengan kasus data sampel yang mengandung pengamatan tidak terdeteksi untuk menggunakan uji parametrik terutama untuk ukuran sampel besar.

b) Program yang dibuat penulis disarankan untuk dapat digunakan oleh peneliti dan praktisi dalam memudahkan menggunakan uji parametrik.

c) Disarankan bagi peneliti lain untuk membandingkan metode uji permutasi yang sifatnya nonparametrik dengan uji parametrik yang diusulkan dalam penelitian ini.

DAFTAR PUSTAKA

Gilbert, R.O. (1987). *Statistical Method for Environmental Pollution Monitoring*. Wiley, New York.

Gleit, A. (1985). Estimation for small normal data sets with detection limits. *Environmental Science and Technology*, Vol. 19, 1201-1206.

Klugman, S. A., Panjer, H. H., dan Willmot, G. E. (2008). *Loss Models: From Data to Decisions*. Edisi ketiga, Wiley, New York.

Mutaqin, A.K. (2015). Kinerja Metode Pengujian Dua Populasi Berdistribusi Log-logistik yang Mengandung Pengamatan Tidak Terdeteksi. *Statistika: Forum Teori dan Aplikasi Statistika*, Vol. 15, No. 1, 31-37.

Mutaqin, A.K., Kudus, A. (2014). Perbandingan Dua Populasi Berdistribusi Log-logistik untuk Data yang Mengandung Pengamatan Tidak Terdeteksi. *Prosiding SnaPP 2014, Sains, Teknologi dan Kesehatan*, 89-94.

Mutaqin, A.K., Kudus, A., Safitri, F.T. (2013). Pendugaan Parameter Distribusi Log-Logistik untuk Data yang Mengandung Pengamatan Tidak Terdeteksi. *Prosiding Seminar Nasional Teknik Industri*, Universitas Malikussaleh, 28-29 Agustus 2013.

Shumway, R. H., Azari, R. S. dan Kayhanian M. (2002). Statistical approaches to estimating mean water quality concentrations with detection limits. *Environ. Sci. Technol*, 36, 3345-3353.

Stoline, M.R. (1993). Comparison of Two Medians Using A Two-Sample Lognormal Model in Environmental Contexts. *Environmetrics*, Vol. 4, No. 3, 323-339.

Warsono. (1996). Analysis of Environmental Pollutant Data Using Generalized Log-logistic Distribution. *Dissertation at University of Alabama at Birmingham*.

Zhong, W., Shukla, R., Succop, P. Levin, L., Welge, J., dan Sivaganesan, S. (2005). Statistical Approaches to Analyze Censored Data with Multiple Detection Limits. *Disertasi Program Doctor of Philosophy University of Cincinnati*.

LAMPIRAN

Lampiran 1. Program Matlab untuk Uji Hipotesis 1 sampai 4

```
function pv1=Test_1_EM(X1,X2,L1,L2)
% Test 1: H1 vs H4
% Input : X1, data terobservasi sampel 1
%         X2, data terobservasi sampel 2
%         L1, data non-detected sampel 1
%         L2, data non-detected sampel 2

% Nilai Log-likehood untuk H1
LL1=LL_H1_EM(X1,X2,L1,L2);

% Nilai Log-likehood untuk H4
LL4=LL_H4_EM(X1,X2,L1,L2);

% Nilai P-Value
CHIS=-2*(LL1-LL4);
pv1=1-chi2cdf(CHIS,2);

function pv2=Test_2_EM(X1,X2,L1,L2)
% Test 1: H2 vs H4
% Input : X1, data terobservasi sampel 1
%         X2, data terobservasi sampel 2
%         L1, data non-detected sampel 1
%         L2, data non-detected sampel 2

% Nilai Log-likehood untuk H2
LL2=LL_H2_EM(X1,X2,L1,L2);

% Nilai Log-likehood untuk H4
LL4=LL_H4_EM(X1,X2,L1,L2);

% Nilai P-Value
CHIS=-2*(LL2-LL4);
pv2=1-chi2cdf(CHIS,2);

function pv3=Test_3_EM(X1,X2,L1,L2)
% Test 1: H1 vs H2
% Input : X1, data terobservasi sampel 1
%         X2, data terobservasi sampel 2
%         L1, data non-detected sampel 1
%         L2, data non-detected sampel 2

% Nilai Log-likehood untuk H1
LL1=LL_H1_EM(X1,X2,L1,L2);

% Nilai Log-likehood untuk H2
LL2=LL_H2_EM(X1,X2,L1,L2);

% Nilai P-Value
CHIS=-2*(LL1-LL2);
pv3=1-chi2cdf(CHIS,2);
```

```
function pv4=Test_4_EM(X1,X2,L1,L2)
% Test 1: H3 vs H4
% Input : X1, data terobservasi sampel 1
%         X2, data terobservasi sampel 2
%         L1, data non-detected sampel 1
%         L2, data non-detected sampel 2

% Nilai Log-likehood untuk H3
LL3=LL_H3_EM(X1,X2,L1,L2);

% Nilai Log-likehood untuk H4
LL4=LL_H4_EM(X1,X2,L1,L2);

% Nilai P-Value
CHIS=-2*(LL3-LL4);
pv4=1-chi2cdf(CHIS,2);

function LL1=LL_H1_EM(X1,X2,L1,L2)
% Program menaksir parameter dan fungsi kemungkinan H1
% Input,  X1: data sampel 1 yang terobservasi
% Input,  X2: data sampel 2 yang terobservasi
% Input,  L1: batas deteksi sampel 1
% Input,  L2: batas deteksi sampel 2
% Output,OUT: taksiran parameter theta dan gamma
% Output, LL: taksiran Log-Likelihood

X1G=[X1;L1];
X2G=[X2;L2];

% % Nilai awal
p251=prctile(X1G,25);
p751=prctile(X1G,75);
g1=2*log(3)/(log(p751)-log(p251));
t1=exp((log(p751)+log(p251))/2);

p252=prctile(X2G,25);
p752=prctile(X2G,75);
g2=2*log(3)/(log(p752)-log(p252));
t2=exp((log(p752)+log(p252))/2);

OUT(1,1)=(t1+t2)/2;
OUT(1,2)=(g1+g2)/2;

delta=10^(-5);
diffe=1;
it=1;

while diffe>=delta
    x0(1)=OUT(it,1);
    x0(2)=OUT(it,2);

    % Tahap-E (Ekspektasi)
    [E11,E12]=EXPECT1(x0(1),x0(2),L1,L2);
    [E21,E22]=EXPECT2(x0(1),x0(2),L1,L2);

    % Optimisasi
```

```
    f = @(x)H1_Fungsi_EM(x,X1,X2,L1,L2,E11,E12);
    x = fsolve(f,x0);
    OUT(it+1,1)=x(1);
    OUT(it+1,2)=x(2);

    % Stopping Rule
    diffe=sqrt(sum((OUT(it+1,1:2)-OUT(it,1:2)).^2));

    % Kenaikan iterasi
    it=it+1;
end
C1=log(x(2))+(x(2)-1)*log(X1)-x(2)*log(x(1))-
2*log(1+(X1/x(1)).^(x(2)));
C2=log(x(2))+(x(2)-1)*log(X2)-x(2)*log(x(1))-
2*log(1+(X2/x(1)).^(x(2)));
C3=x(2)*E11-x(2)*log(x(1))-E21;
C4=x(2)*E12-x(2)*log(x(1))-E22;
LL1=sum(C1)+sum(C2)+sum(C3)+sum(C4);


function LL2=LL_H2_EM(X1,X2,L1,L2)
% Program menaksir parameter dan fungsi kemungkinan H2
% Input,  X1: data sampel 1 yang terobservasi
% Input,  X2: data sampel 2 yang terobservasi
% Input,  L1: batas deteksi sampel 1
% Input,  L2: batas deteksi sampel 2
% Output,LL2: taksiran Log-Likelihood H2

% % Nilai awal
p251=prctile(X1,25);
p751=prctile(X1,75);
g1=2*log(3)/(log(p751)-log(p251));
t1=exp((log(p751)+log(p251))/2);

p252=prctile(X2,25);
p752=prctile(X2,75);
g2=2*log(3)/(log(p752)-log(p252));
t2=exp((log(p752)+log(p252))/2);

OUT(1,1)=t1;
OUT(1,2)=t2;
OUT(1,3)=(g1+g2)/2;

delta=10^(-5);
diffe=1;
it=1;

while diffe>=delta
    x0(1)=OUT(it,1);
    x0(2)=OUT(it,2);
    x0(3)=OUT(it,3);

    % Tahap-E (Ekspektasi)
    [E31,E32]=EXPECT3(x0(1),x0(2),x0(3),L1,L2);
    [E41,E42]=EXPECT4(x0(1),x0(2),x0(3),L1,L2);

    % Optimisasi
```

```
    f = @(x)H2_Fungsi_EM(x,X1,X2,L1,L2,E31,E32);
    x = fsolve(f,x0);
    OUT(it+1,1)=x(1);
    OUT(it+1,2)=x(2);
    OUT(it+1,3)=x(3);

    % Stopping Rule
    diffe=sqrt(sum((OUT(it+1,1:3)-OUT(it,1:3)).^2));

    % Kenaikan iterasi
    it=it+1;
end
C1=log(x(3))+(x(3)-1)*log(X1)-x(3)*log(x(1))-
2*log(1+(X1/x(1)).^(x(3)));
C2=log(x(3))+(x(3)-1)*log(X2)-x(3)*log(x(2))-
2*log(1+(X2/x(2)).^(x(3)));
C3=x(3)*E31-x(3)*log(x(1))-E41;
C4=x(3)*E32-x(3)*log(x(2))-E42;
LL2=sum(C1)+sum(C2)+sum(C3)+sum(C4);


function LL3=LL_H3_EM(X1,X2,L1,L2)
% Program menaksir parameter dan fungsi kemungkinan H3
% Input,  X1: data sampel 1 yang terobservasi
% Input,  X2: data sampel 2 yang terobservasi
% Input,  L1: batas deteksi sampel 1
% Input,  L2: batas deteksi sampel 2
% Output,LL3: taksiran Log-Likelihood

% % Nilai awal
p251=prctile(X1,25);
p751=prctile(X1,75);
g1=2*log(3)/(log(p751)-log(p251));
t1=exp((log(p751)+log(p251))/2);

p252=prctile(X2,25);
p752=prctile(X2,75);
g2=2*log(3)/(log(p752)-log(p252));
t2=exp((log(p752)+log(p252))/2);

OUT(1,1)=(t1+t2)/2;
OUT(1,2)=g1;
OUT(1,3)=g2;

delta=10^(-5);
diffe=1;
it=1;

while diffe>=delta
    x0(1)=OUT(it,1);
    x0(2)=OUT(it,2);
    x0(3)=OUT(it,3);

    % Tahap-E (Ekspektasi)
    [E51,E52]=EXPECT5(x0(1),x0(2),x0(3),L1,L2);
    [E61,E62]=EXPECT6(x0(1),x0(2),x0(3),L1,L2);
```

```
    % Optimisasi
    f = @(x)H3_Fungsi_EM(x,X1,X2,L1,L2,E51,E52);
    x = fsolve(f,x0);
    OUT(it+1,1)=x(1);
    OUT(it+1,2)=x(2);
    OUT(it+1,3)=x(3);

    % Stopping Rule
    diffe=sqrt(sum((OUT(it+1,1:3)-OUT(it,1:3)).^2));

    % Kenaikan iterasi
    it=it+1;
end
C1=log(x(2))+(x(2)-1)*log(X1)-x(2)*log(x(1))-
2*log(1+(X1/x(1)).^(x(2)));
C2=log(x(3))+(x(3)-1)*log(X2)-x(3)*log(x(1))-
2*log(1+(X2/x(1)).^(x(3)));
C3=x(2)*E51-x(2)*log(x(1))-E61;
C4=x(3)*E52-x(3)*log(x(1))-E62;
LL3=sum(C1)+sum(C2)+sum(C3)+sum(C4);


function LL4=LL_H4_EM(X1,X2,L1,L2)
% Program menaksir parameter dan fungsi kemungkinan H4
% Input,  X1: data sampel 1 yang terobservasi
% Input,  X2: data sampel 2 yang terobservasi
% Input,  L1: batas deteksi sampel 1
% Input,  L2: batas deteksi sampel 2
% Output,LL4: taksiran Log-Likelihood H4

% % Nilai awal
p251=prctile(X1,25);
p751=prctile(X1,75);
g1=2*log(3)/(log(p751)-log(p251));
t1=exp((log(p751)+log(p251))/2);

p252=prctile(X2,25);
p752=prctile(X2,75);
g2=2*log(3)/(log(p752)-log(p252));
t2=exp((log(p752)+log(p252))/2);

OUT(1,1)=t1;
OUT(1,2)=t2;
OUT(1,3)=g1;
OUT(1,4)=g2;

delta=10^(-5);
diffe=1;
it=1;

while diffe>=delta
    x0(1)=OUT(it,1);
    x0(2)=OUT(it,2);
    x0(3)=OUT(it,3);
    x0(4)=OUT(it,4);

    % Tahap-E (Ekspektasi)
```

```
    [E71,E72]=EXPECT7(x0(1),x0(2),x0(3),x0(4),L1,L2);
    [E81,E82]=EXPECT8(x0(1),x0(2),x0(3),x0(4),L1,L2);

    % Optimisasi
    f = @(x)H4_Fungsi_EM(x,X1,X2,L1,L2,E71,E72);
    x = fsolve(f,x0);
    OUT(it+1,1)=x(1);
    OUT(it+1,2)=x(2);
    OUT(it+1,3)=x(3);
    OUT(it+1,4)=x(4);

    % Stopping Rule
    diffe=sqrt(sum((OUT(it+1,1:4)-OUT(it,1:4)).^2));

    % Kenaikan iterasi
    it=it+1;
end
C1=log(x(3))+(x(3)-1)*log(X1)-x(3)*log(x(1))-
2*log(1+(X1/x(1)).^(x(3)));
C2=log(x(4))+(x(4)-1)*log(X2)-x(4)*log(x(2))-
2*log(1+(X2/x(2)).^(x(4)));
C3=x(3)*E71-x(3)*log(x(1))-E81;
C4=x(4)*E72-x(4)*log(x(2))-E82;
LL4=sum(C1)+sum(C2)+sum(C3)+sum(C4);


function F = H1_Fungsi_EM(x,X1,X2,L1,L2,E11,E12)

r1=length(X1);
r2=length(X2);
n1=r1+length(L1);
n2=r2+length(L2);

% Turunan 1 terhadap theta

C1=2*(x(2)/x(1))*exp(x(2)*log(X1/x(1)))./(1+exp(x(2)*log(X1/x(1))));

C2=2*(x(2)/x(1))*exp(x(2)*log(X2/x(1)))./(1+exp(x(2)*log(X2/x(1))));
    T11=-(n1+n2)*x(2)/x(1)+sum(C1)+sum(C2);

    % Turunan 1 terhadap gamma
    C3=log(X1)-
2*log(X1/x(1)).*exp(x(2)*log(X1/x(1)))./(1+exp(x(2)*log(X1/x(1))));
    C4=log(X2)-
2*log(X2/x(1)).*exp(x(2)*log(X2/x(1)))./(1+exp(x(2)*log(X2/x(1))));
    T12=(r1+r2)/x(2)-
(n1+n2)*log(x(1))+sum(C3)+sum(C4)+sum(E11)+sum(E12);

    F = [T11;T12];

function F = H2_Fungsi_EM(x,X1,X2,L1,L2,E31,E32)

r1=length(X1);
r2=length(X2);
n1=r1+length(L1);
n2=r2+length(L2);
```

```
    % Turunan 1 terhadap theta-1

C1=2*(x(3)/x(1))*exp(x(3)*log(X1/x(1)))./(1+exp(x(3)*log(X1/x(1))));

C2=2*(x(3)/x(2))*exp(x(3)*log(X2/x(2)))./(1+exp(x(3)*log(X2/x(2))));
    T11=-n1*x(3)/x(1)+sum(C1);
    T12=-n2*x(3)/x(2)+sum(C2);

    % Turunan 1 terhadap gamma
    C3=log(X1)-
2*log(X1/x(1)).*exp(x(3)*log(X1/x(1)))./(1+exp(x(3)*log(X1/x(1))));
    C4=log(X2)-
2*log(X2/x(2)).*exp(x(3)*log(X2/x(2)))./(1+exp(x(3)*log(X2/x(2))));
    T13=(r1+r2)/x(3)-
(n1*log(x(1))+n2*log(x(2)))+sum(C3)+sum(C4)+sum(E31)+sum(E32);

    F = [T11;T12;T13];

function F = H3_Fungsi_EM(x,X1,X2,L1,L2,E51,E52)

r1=length(X1);
r2=length(X2);
n1=r1+length(L1);
n2=r2+length(L2);

% Turunan 1 terhadap theta

C1=2*(x(2)/x(1))*exp(x(2)*log(X1/x(1)))./(1+exp(x(2)*log(X1/x(1))));

C2=2*(x(3)/x(1))*exp(x(3)*log(X2/x(1)))./(1+exp(x(3)*log(X2/x(1))));
    T11=-(n1*x(2)+n2*x(3))/x(1)+sum(C1)+sum(C2);

    % Turunan 1 terhadap gamma
    C3=log(X1)-
2*log(X1/x(1)).*exp(x(2)*log(X1/x(1)))./(1+exp(x(2)*log(X1/x(1))));
    C4=log(X2)-
2*log(X2/x(1)).*exp(x(3)*log(X2/x(1)))./(1+exp(x(3)*log(X2/x(1))));
    T12=r1/x(2)-n1*log(x(1))+sum(C3)+sum(E51);
    T13=r2/x(3)-n2*log(x(1))+sum(C4)+sum(E52);

    F = [T11;T12;T13];

function F = H4_Fungsi_EM(x,X1,X2,L1,L2,E71,E72)

r1=length(X1);
r2=length(X2);
n1=r1+length(L1);
n2=r2+length(L2);

    % Turunan 1 terhadap theta

C1=2*(x(3)/x(1))*exp(x(3)*log(X1/x(1)))./(1+exp(x(3)*log(X1/x(1))));

C2=2*(x(4)/x(2))*exp(x(4)*log(X2/x(2)))./(1+exp(x(4)*log(X2/x(2))));
    T11=-n1*x(3)/x(1)+sum(C1);
    T12=-n2*x(4)/x(2)+sum(C2);
```

```
    % Turunan 1 terhadap gamma
    C3=log(X1)-
2*log(X1/x(1)).*exp(x(3)*log(X1/x(1)))./(1+exp(x(3)*log(X1/x(1))));
    C4=log(X2)-
2*log(X2/x(2)).*exp(x(4)*log(X2/x(2)))./(1+exp(x(4)*log(X2/x(2))));
    T13=r1/x(3)-n1*log(x(1))+sum(C3)+sum(E71);
    T14=r2/x(4)-n2*log(x(2))+sum(C4)+sum(E72);

    F = [T11;T12;T13;T14];

function [E11,E12]=EXPECT1(t,g,L1,L2)
% t     : parameter theta
% g     : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E11: tahap ekspektasi E1 utk sampel 1
% Output, E12: tahap ekspektasi E1 utk sampel 2

F11=((L1/t).^g./(1+(L1/t).^g));
E11=log(t)+log(F11)/g+(1./F11-1).*log(1-F11)/g;

F12=((L2/t).^g./(1+(L2/t).^g));
E12=log(t)+log(F12)/g+(1./F12-1).*log(1-F12)/g;


function [E21,E22]=EXPECT2(t,g,L1,L2)
% t     : parameter theta
% g     : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E21: tahap ekspektasi E2 utk sampel 1
% Output, E22: tahap ekspektasi E2 utk sampel 2

F21=((L1/t).^g./(1+(L1/t).^g));
E21=log(1+F21./(1-F21))+log(1-F21)./F21+1;

F22=((L2/t).^g./(1+(L2/t).^g));
E22=log(1+F22./(1-F22))+log(1-F22)./F22+1;


function [E31,E32]=EXPECT3(t1,t2,g,L1,L2)
% t1,t2 : parameter theta
% g     : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E31: tahap ekspektasi E3 utk sampel 1
% Output, E32: tahap ekspektasi E3 utk sampel 2

F31=((L1/t1).^g./(1+(L1/t1).^g));
E31=log(t1)+log(F31)/g+(1./F31-1).*log(1-F31)/g;

F32=((L2/t2).^g./(1+(L2/t2).^g));
E32=log(t2)+log(F32)/g+(1./F32-1).*log(1-F32)/g;


function [E41,E42]=EXPECT4(t1,t2,g,L1,L2)
% t1,t2 : parameter theta
% g     : parameter gamma
```

```
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E41: tahap ekspektasi E4 utk sampel 1
% Output, E42: tahap ekspektasi E4 utk sampel 2

F41=((L1/t1).^g./(1+(L1/t1).^g));
E41=log(1+F41./(1-F41))+log(1-F41)./F41+1;

F42=((L2/t2).^g./(1+(L2/t2).^g));
E42=log(1+F42./(1-F42))+log(1-F42)./F42+1;


function [E51,E52]=EXPECT5(t,g1,g2,L1,L2)
% t     : parameter theta
% g1,g2 : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E51: tahap ekspektasi E5 utk sampel 1
% Output, E52: tahap ekspektasi E5 utk sampel 2

F51=((L1/t).^g1./(1+(L1/t).^g1));
E51=log(t)+log(F51)/g1+(1./F51-1).*log(1-F51)/g1;

F52=((L2/t).^g2./(1+(L2/t).^g2));
E52=log(t)+log(F52)/g2+(1./F52-1).*log(1-F52)/g2;


function [E61,E62]=EXPECT6(t,g1,g2,L1,L2)
% t     : parameter theta
% g1,g2 : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E61: tahap ekspektasi E6 utk sampel 1
% Output, E62: tahap ekspektasi E6 utk sampel 2

F61=((L1/t).^g1./(1+(L1/t).^g1));
E61=log(1+F61./(1-F61))+log(1-F61)./F61+1;

F62=((L2/t).^g2./(1+(L2/t).^g2));
E62=log(1+F62./(1-F62))+log(1-F62)./F62+1;


function [E71,E72]=EXPECT7(t1,t2,g1,g2,L1,L2)
% t1,t2 : parameter theta
% g1,g2 : parameter gamma
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E71: tahap ekspektasi E7 utk sampel 1
% Output, E72: tahap ekspektasi E7 utk sampel 2

F71=((L1/t1).^g1./(1+(L1/t1).^g1));
E71=log(t1)+log(F71)/g1+(1./F71-1).*log(1-F71)/g1;

F72=((L2/t2).^g2./(1+(L2/t2).^g2));
E72=log(t2)+log(F72)/g2+(1./F72-1).*log(1-F72)/g2;


function [E81,E82]=EXPECT8(t1,t2,g1,g2,L1,L2)
% t1,t2 : parameter theta
% g1,g2 : parameter gamma
```

```
% L1,L2 : batas deteksi sampel 1 dan 2
% Output, E81: tahap ekspektasi E8 utk sampel 1
% Output, E82: tahap ekspektasi E8 utk sampel 2

F81=((L1/t1).^g1./(1+(L1/t1).^g1));
E81=log(1+F81./(1-F81))+log(1-F81)./F81+1;

F82=((L2/t2).^g2./(1+(L2/t2).^g2));
E82=log(1+F82./(1-F82))+log(1-F82)./F82+1;
```

Lampiran 2. Program untuk Simulasi Monte Carlo

```
function KT1=SIM123_EM(N,p,a1,b1,a2,b2)
% Fungsi untuk menguji Hipotesis H1, H2, H3
alpha=0.05;
SIM=5;
for i=1:SIM
    % Sampel 1
    [X1,BD1,t1]=SkemaA1(N,a1,b1,p);
    L1=BD1*ones(t1,1);
    % Sampel 2
    [X2,BD2,t2]=SkemaA1(N,a2,b2,p);
    L2=BD2*ones(t2,1);

    % OUT(i,1)=Test_1_EM(X1,X2,L1,L2);
    % OUT(i,1)=Test_2_EM(X1,X2,L1,L2);
    % OUT(i,1)=Test_3_EM(X1,X2,L1,L2);
    OUT(i,1)=Test_4_EM(X1,X2,L1,L2);

end
IND=OUT<=alpha;
KT1=sum(IND)/SIM;


function [X,BD,t]=SkemaA1(N,a,b,p)
% Program untuk membangkitkan data dengan 1 BD
% N adalah sample size
% a dan b adalah parameter distribusi LL
% p adalah persentase data yang tersensor

% Banyaknya BD
kel=1;

% Pembangkitan Data Distribusi LL
A = GenerateLL(N,a,b);
B = rand(N,1);
C = [A,B];
D = sortrows(C,2);
baris=N/kel;
MAT=reshape(D(:,1),baris,kel);
datasort=sort(MAT);
sensor=floor(N/kel*p)+1;
t=sensor-1;
% Pemilihan BD
for i=1:kel
    BD(i)=datasort(sensor(i),i);
end
X1=datasort(sensor(1):baris,1);
X=[X1];

function X = GenerateLL(n,a,b)
% a = alpha dalam model TLL
% b = beta dalam model TLL
% program membangkitkan data berukuran n
% berdistribusi Log-Logistic, LL(a,b)
% rand('state',100)
for i=1:n
```

```
    U=rand(1,1);
    X(i,1)=(U/((1-U)*exp(b)))^(1/a);
end
```

Perbandingan Median Populasi Log-Logistik untuk Data yang Mengandung Pengamatan Tidak Terdeteksi

Aceng Komarudin Mutaqin, Siti Sunendiari
Program Studi Statistika, Universitas Islam Bandung
Jl. Ranggamalela No. 1 Bandung 40116
Email: aceng.k.mutaqin@gmail.com

Abstrak

Suatu prosedur diturunkan untuk membandingkan dua median populasi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi. Diasumsikan bahwa dua sampel saling bebas berukuran n_1 dan n_2 berasal dari dua populasi log-logistik $LLD(\theta_1, \gamma_1)$ dan $LLD(\theta_2, \gamma_2)$. Penaksiran kemungkinan maksimum digunakan untuk menaksir parameter dalam empat kasus: $H_1(\theta_1 = \theta_2, \gamma_1 = \gamma_2)$, $H_2(\theta_1 \neq \theta_2, \gamma_1 = \gamma_2)$, $H_3(\theta_1 = \theta_2, \gamma_1 \neq \gamma_2)$, dan $H_1(\theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2)$. Suatu pedoman digambarkan untuk menguji kesamaan dua median ($H_0: \theta_1 = \theta_2$ lawan $H_r: \theta_1 \neq \theta_2$). Dua prosedur diusulkan untuk uji di atas, tergantung apakah koefisien variasinya sama: ($H_0: \gamma_1 = \gamma_2$) atau tidak ($H_0: \gamma_1 \neq \gamma_2$). Uji chi-square asimtotik digunakan dalam uji di atas. Dalam makalah ini diberikan suatu contoh kasus menggunakan data emisi gas buang kendaraan bermotor.

Keywords: log-logistic, data tidak terdeteksi, penaksiran kemungkinan maksimum, koefisien variasi, emisi gas buang kendaraan.

1. Pendahuluan

Nilai pengamatan yang tidak terdeteksi menjadi suatu masalah yang sulit ketika tujuannya adalah membandingkan dua populasi yang berbeda. Secara umum ada dua pendekatan yang diusulkan untuk permasalahan tersebut, yaitu pendekatan parametrik dan nonparametrik. Untuk dua data sampel dari dua populasi yang berbeda mengikuti distribusi lognormal, Stoline (1993) mengusulkan menggunakan uji kesamaan dua median untuk membandingkan dua populasi ketika data mengandung pengamatan tidak terdeteksi. Sementara itu, Zhong dkk. (2005) menggunakan informasi fungsi kemungkinan untuk pengujiannya. Untuk kasus yang sama, uji standar seperti uji T seringkali digunakan oleh para peneliti (Zhong dkk., 2005).

Selain distribusi lognormal, distribusi lain yang bisa digunakan untuk memodelkan data lingkungan adalah distribusi log-logistik (Warsono, 1996). Mutaqin dan Kudus (2014) membahas uji permutasi untuk membandingkan dua populasi berdistribusi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi dengan jalan membandingkan kedua median populasinya. Hasil penelitian Mutaqin (2015) menunjukkan bahwa untuk data yang mengandung batas deteksi tunggal, uji permutasi tersebut akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi rendah dan perbedaan besar koefisien variasinya. Sedangkan untuk data yang mengandung batas deteksi ganda, uji permutasi tersebut akan baik digunakan untuk kasus ukuran sampel besar, persentase pengamatan tidak terdeteksi kedua sampel berbeda dan perbedaan besar koefisien variasi. Dalam makalah ini diusulkan metode yang sifatnya parametrik untuk menguji kesamaan median dua populasi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi menggunakan informasi fungsi kemungkinan.

Suatu contoh kasus diberikan untuk menerapkan metode di atas menggunakan data emisi gas buang kendaraan bermotor.

2. Distribusi Log-Logistik

Fungsi densitas dari distribusi log-logistik dengan parameter $\alpha > 0$ dan parameter lokasi $-\infty < \beta < \infty$ adalah

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

Mutaqin dkk. (2013) membahas metode penaksiran kemungkinan maksimum untuk parameter α, dan β yang data sampelnya mengandung pengamatan tidak terdeteksi menggunakan algoritme EM. Melalui parameterisasi ulang, dengan memisalkan $\gamma = \alpha$, dan $\theta = e^{-\beta/\alpha}$, akan diperoleh fungsi densitas dari distribusi log-logistiknya adalah

$$f(x;\theta,\gamma) = \frac{\gamma(x/\theta)^{\gamma}}{x[1+(x/\theta)^{\gamma}]^2}; x > 0,$$

dimana $\gamma > 0$ adalah parameter bentuk, dan $\theta > 0$ adalah parameter skala (Klugman dkk., 2012). Dapat ditunjukkan bahwa median dari distribusi log-logistik di atas adalah $M = \theta$, sedangkan koefisien variasinya adalah

$$CV = \frac{(E[X^2]-(E[X])^2)^{1/2}}{E[X]} = \frac{\left[\Gamma\left(1+\frac{2}{\gamma}\right)\Gamma\left(1-\frac{2}{\gamma}\right)-\left[\Gamma\left(1+\frac{1}{\gamma}\right)\Gamma\left(1-\frac{1}{\gamma}\right)\right]^2\right]^{1/2}}{\Gamma\left(1+\frac{1}{\gamma}\right)\Gamma\left(1-\frac{1}{\gamma}\right)}.$$

Terlihat bahwa median merupakan fungsi dari parameter θ, sedangkan koefisien variasinya merupakan fungsi dari parameter γ.

3. Uji Permutasi untuk Kesamaan Dua Median Distribusi Log-logistik

Misalkan θ_1, γ_1 dan θ_2, γ_2 masing-masing menyatakan parameter dari dua populasi berdistribusi log-logistik. Misalkan juga bahwa median kedua populasi tersebut adalah M_1 dan M_2. Kedua median dinyatakan sama ketika hipotesis $H_0: \theta_1 = \theta_2$ diterima. Untuk kasus homogen ($\gamma_1 = \gamma_2$), rata-rata dan varians dari distribusi log-logistik untuk kedua populasi mungkin saja berbeda, tetapi koefisien variasinya sama. Jika hipotesis nol diterima dalam kasus homogen, maka dapat disimpulkan bahwa kedua populasi identik. Untuk kasus heterogen ($\gamma_1 \neq \gamma_2$), jika hipotesis nol diterima, maka hanya dapat disimpulkan bahwa kedua median populasi identik.

Mutaqin dan Kudus (2014) membahas pengujian hipotesis $H_0: \theta_1 = \theta_2$ melawan $H_1: \theta_1 \neq \theta_2$ menggunakan uji permutasi. Misalkan $x_{11}, \cdots, x_{1n_1}$ dan $x_{21}, \cdots, x_{2n_2}$ masing-masing menyatakan sampel-sampel saling bebas yang berukuran n_1 dan n_2 dari dua populasi log-logistik, $LLD(\theta_1, \gamma_1)$ dan $LLD(\theta_2, \gamma_2)$. Diasumsikan bahwa untuk setiap x_{ij} ada batas deteksi L_{ij}, untuk $i = 1,2$ dan $j = 1,2,\cdots,n_i$. Jika nilai pengamatannya terdeteksi, maka x_{ij} yang dicatat. Sedangkan jika nilai pengamatannya tidak terdeteksi ($< L_{ij}$), maka L_{ij} yang dicatat (tersensor kiri). Misalkan untuk sampel i ada r_i pengamatan yang terdeteksi, sisanya $n_i - r_i$ pengamatan tidak terdeteksi. Tahapan yang rinci dari uji permutasi untuk hipotesis $H_0: \theta_1 = \theta_2$ melawan $H_1: \theta_1 \neq \theta_2$ dapat dilihat di Mutaqin dan Kudus (2014).

4. Prosedur Pengujian yang Diusulkan

Dengan mengikuti hasil dari Stoline (1993), definisikan empat hipotesis berikut:

$$H_1: \theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma;$$

$$H_2: \theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma;$$

$H_3: \theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2$;

$H_4: \theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$.

Empat pengujian hipotesis akan diusulkan, yaitu

uji hipotesis 1: H_1 lawan H_4; (1)

uji hipotesis 2: H_2 lawan H_4; (2)

uji hipotesis 3: H_1 lawan H_2; (3)

uji hipotesis 4: H_3 lawan H_4. (4)

Berikut ini akan direkomendasikan strategi pengujian melalui tiga tahapan, yaitu:

Tahap 1: menguji kehomogenan secara keseluruhan untuk dua populasi log-logistik yang data sampelnya mengandung pengamatan tidak terdeteksi menggunakan uji hipotesis 1.

Tahap 2: menguji kehomogenan koefisien variasi menggunakan uji hipotesis 2.

Tahap 3: menguji kesamaan dua median menggunakan uji hipotesis 3 untuk kasus koefisien variasi homogen, atau menggunakan uji hipotesis 4 untuk kasus koefisien variasi tidak homogen.

Uji hipotesis 1 sampai 4 dilakukan dengan menggunakan informasi nilai fungsi kemungkinan dari hipotesis H_1, H_2, H_3 dan H_4. Berdasarkan nilai-nilai fungsi kemungkinan tersebut, dibentuk statistik uji chi-kuadrat asimtotik untuk semua uji hipotesis 1 sampai 4. Fungsi log-kemungkinan untuk kasus hipotesis $H_1: \theta_1 = \theta_2 = \theta, \gamma_1 = \gamma_2 = \gamma$ adalah

$$l_1 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta - \ln\left[1+\left(L_{ij}/\theta\right)^{\gamma}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_1 terhadap parameter θ dan γ masing-masing adalah

$$\frac{\partial l_1}{\partial\theta} = -(n_1+n_2)\frac{\gamma}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma/\theta)\left(x_{ij}/\theta\right)^{\gamma}}{1+\left(x_{ij}/\theta\right)^{\gamma}} + \sum_{j=r_i+1}^{n_i}\frac{(\gamma/\theta)\left(L_{ij}/\theta\right)^{\gamma}}{1+\left(L_{ij}/\theta\right)^{\gamma}}\right] \tag{5}$$

dan

$$\frac{\partial l_1}{\partial\gamma} = \frac{(r_1+r_2)}{\gamma} - (n_1+n_2)\ln\theta + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(x_{ij}/\theta\right)\left(x_{ij}/\theta\right)^{\gamma}}{1+\left(x_{ij}/\theta\right)^{\gamma}}\right\} + \sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(L_{ij}/\theta\right)\left(L_{ij}/\theta\right)^{\gamma}}{1+\left(L_{ij}/\theta\right)^{\gamma}}\right\}\right]. \tag{6}$$

Fungsi log-kemungkinan untuk kasus hipotesis $H_2: \theta_1 \neq \theta_2, \gamma_1 = \gamma_2 = \gamma$ adalah

$$l_2 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma - \gamma\ln\theta_i + (\gamma-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}\right]\right\}\right.$$
$$\left.+\sum_{j=r_i+1}^{n_i}\left\{\gamma\ln L_{ij} - \gamma\ln\theta_i - \ln\left[1+\left(L_{ij}/\theta_i\right)^{\gamma}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_2 terhadap parameter θ_1, θ_2 dan γ masing-masing adalah

$$\frac{\partial l_2}{\partial\theta_i} = -n_i\frac{\gamma}{\theta_i} + \sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}}{1+\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma}} + \sum_{j=r_i+1}^{n_i}\frac{(\gamma/\theta_i)\left(L_{ij}/\theta_i\right)^{\gamma}}{1+\left(L_{ij}/\theta_i\right)^{\gamma}}; i = 1,2 \qquad (7)$$

dan

$$\frac{\partial l_2}{\partial\gamma} = \frac{(r_1+r_2)}{\gamma} - (n_1\ln\theta_1 + n_2\ln\theta_2)$$
$$+\sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(x_{ij}/\theta_i\right)\left(x_{ij}/\theta_i\right)^{\gamma}}{1+\left(x_{ij}/\theta_i\right)^{\gamma}}\right\}\right. \qquad (8)$$
$$\left.+\sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(L_{ij}/\theta_i\right)\left(L_{ij}/\theta_i\right)^{\gamma}}{1+\left(L_{ij}/\theta_i\right)^{\gamma}}\right\}\right].$$

Fungsi log-kemungkinan untuk kasus hipotesis $H_3: \theta_1 = \theta_2 = \theta, \gamma_1 \neq \gamma_2$ adalah

$$l_3 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i\ln\theta + (\gamma_i-1)\ln x_{ij} - 2\ln\left[1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}\right]\right\}\right.$$
$$\left.+\sum_{j=r_i+1}^{n_i}\left\{\gamma_i\ln L_{ij} - \gamma_i\ln\theta - \ln\left[1+\left(L_{ij}/\theta\right)^{\gamma_i}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_3 terhadap parameter θ, γ_1, dan γ_2 masing-masing adalah

$$\frac{\partial l_3}{\partial\theta} = -\frac{(n_1\gamma_1+n_2\gamma_2)}{\theta} + \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\frac{2(\gamma_i/\theta)\left(x_{ij}/\theta\right)^{\gamma_i}}{1+\left(x_{ij}/\theta\right)^{\gamma_i}} + \sum_{j=r_i+1}^{n_i}\frac{(\gamma_i/\theta)\left(L_{ij}/\theta\right)^{\gamma_i}}{1+\left(L_{ij}/\theta\right)^{\gamma_i}}\right] \qquad (9)$$

dan

$$\frac{\partial l_3}{\partial\gamma_i} = \frac{r_i}{\gamma_i} - n_i\ln\theta + \sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta}\right)\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}{1+\left(\frac{x_{ij}}{\theta}\right)^{\gamma_i}}\right\} + \qquad (10)$$
$$+\sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(L_{ij}/\theta\right)\left(L_{ij}/\theta\right)^{\gamma_i}}{1+\left(L_{ij}/\theta\right)^{\gamma_i}}\right\}; i = 1,2$$

Fungsi log-kemungkinan untuk kasus hipotesis $H_4: \theta_1 \neq \theta_2, \gamma_1 \neq \gamma_2$ adalah

$$l_4 = \sum_{i=1}^{2}\left[\sum_{j=1}^{r_i}\left\{\ln\gamma_i - \gamma_i \ln\theta_i + (\gamma_i - 1)\ln x_{ij} - 2\ln\left[1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}\right]\right\} + \sum_{j=r_i+1}^{n_i}\left\{\gamma_i \ln L_{ij} - \gamma_i \ln\theta_i - \ln\left[1 + \left(L_{ij}/\theta_i\right)^{\gamma_i}\right]\right\}\right]$$

Turunan pertama dari fungsi log-kemungkinan l_4 terhadap parameter θ_1, θ_2, γ_1, dan γ_2 masing-masing adalah

$$\frac{\partial l_4}{\partial\theta_i} = -n_i\frac{\gamma_i}{\theta_i} + \sum_{j=1}^{r_i}\frac{2\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}} + \sum_{j=r_i+1}^{n_i}\frac{\left(\frac{\gamma_i}{\theta_i}\right)\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}{1 + \left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}; i = 1, 2 \tag{11}$$

dan

$$\frac{\partial l_4}{\partial\gamma_i} = \frac{r_i}{\gamma_i} - n_i \ln\theta_i + \sum_{j=1}^{r_i}\left\{\ln x_{ij} - \frac{2\ln\left(\frac{x_{ij}}{\theta_i}\right)\left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}{1 + \left(\frac{x_{ij}}{\theta_i}\right)^{\gamma_i}}\right\} + \sum_{j=r_i+1}^{n_i}\left\{\ln L_{ij} - \frac{\ln\left(\frac{L_{ij}}{\theta_i}\right)\left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}{1 + \left(\frac{L_{ij}}{\theta_i}\right)^{\gamma_i}}\right\}; i = 1, 2. \tag{12}$$

Uji chi-square asimtotik dapat digunakan untuk menguji hipotesis 1 sampai uji hipotesis 4 dengan statistika ujinya masing-masing adalah

$$\chi_1^2 = -2\left(\hat{l}_1 - \hat{l}_4\right) \sim \chi^2_{2,1-\alpha},$$
$$\chi_2^2 = -2\left(\hat{l}_2 - \hat{l}_4\right) \sim \chi^2_{1,1-\alpha},$$
$$\chi_3^2 = -2\left(\hat{l}_1 - \hat{l}_2\right) \sim \chi^2_{1,1-\alpha},$$
$$\chi_4^2 = -2\left(\hat{l}_3 - \hat{l}_4\right) \sim \chi^2_{1,1-\alpha},$$

dimana $\chi^2_{r,1-\alpha}$ adalah kuantil ke-$100(1-\alpha)$ dari distribusi chi-square dengan derajat bebas r. Sementara itu $\hat{l}_1$, $\hat{l}_2$, $\hat{l}_3$, dan $\hat{l}_4$ masing-masing merupakan taksiran kemungkinan maksimum untuk nilai fungsi log-kemungkinan pada hipotesis H_1, H_2, H_3, dan H_4.

5. Contoh Numerik

Dalam bagian ini akan diberikan contoh kasus penerapan metode yang diusulkan pada data kadar carbon monoksida (CO) hasil uji emisi gas buang kendaraan bermotor dua pabrikan kendaraan A dan B di Kota Bandung. Tabel 1 menyajikan data kadar CO dalam persen untuk kendaraan-kendaraan kedua pabrikan tersebut.

Dengan menerapkan metode yang diusulkan diperoleh nilai p-value untuk uji hipotesis 1, uji hipotesis 2 dan uji hipotesis 3 masing-masing sebesar 0,5925; 0,6162; dan 0,9615. Berdasarkan nilai-nilai p-value tersebut dapat disimpulkan bahwa secara statistik median dan distribusi kadar CO untuk kendaraan-kendaraan pabrikan A dan B di Kota Bandung adalah sama.

Tabel 1. Data Kadar CO (dalam %)

Pabrikan *A*	Pabrikan *B*
0,00*	0,00*
0,01	0,00*
0,01	0,01
0,01	0,01
0,02	0,01
0,02	0,01
0,03	0,01
0,05	0,02
0,09	0,02
2,21	0,02
2,29	0,17
	0,18
	0,47
	1,50
	2,53

*tidak terdeteksi (batas deteksi 0,01)

6. Diskusi

Dalam artikel ini telah dibahas suatu metode untuk membandingkan dua median populasi log-logistik untuk data yang mengandung pengamatan tidak terdeteksi. Selain untuk membandingkan dua median, metode yang telah diusulkan juga menghasilkan perbandingan dua koefisien variasi untuk dua populasi log-logistik tersebut serta dapat menjawab pertanyaan apakah dua populasi log-logistik tersebut identik atau tidak.

Dalam artikel ini taksiran parameter distribusi log-logistiknya menggunakan metode penaksiran kemungkinan maksimum dengan memperlakukan pengamatan yang tidak terdeteksinya sebagai data tersensor kiri. Perlu dikembangkan metode penaksiran lain misalnya menggunakan algoritme EM (*expectation maximization*) untuk menaksir parameter distribusi log-logistiknya.

7. Daftar Pustaka

Klugman, S. A., Panjer, H. H., dan Willmot, G. E. (2012). *Loss Models: From Data to Decisions.* Edisi Keempat, Wiley, New York.

Mutaqin, A.K. (2015). Kinerja Metode Pengujian Dua Populasi Berdistribusi Log-logistik yang Mengandung Pengamatan Tidak Terdeteksi. *Statistika: Forum Teori dan Aplikasi Statistika*, Vol. 15, No. 1, 31-37.

Mutaqin, A.K., Kudus, A. (2014). Perbandingan Dua Populasi Berdistribusi Log-logistik untuk Data yang Mengandung Pengamatan Tidak Terdeteksi. *Prosiding SnaPP 2014, Sains, Teknologi dan Kesehatan*, 89-94.

Mutaqin, A.K., Kudus, A., Safitri, F.T. (2013). Pendugaan Parameter Distribusi Log-Logistik untuk Data yang Mengandung Pengamatan Tidak Terdeteksi. *Prosiding Seminar Nasional Teknik Industri*, Universitas Malikussaleh, 28-29 Agustus 2013.

Stoline, M.R. (1993). Comparison of Two Medians Using A Two-Sample Lognormal Model in Environmental Contexts. *Environmetrics*, Vol. 4, No. 3, 323-339.

Warsono. (1996). Analysis of Environmental Pollutant Data Using Generalized Log-logistic Distribution. *Dissertation at University of Alabama at Birmingham.*

Zhong, W., Shukla, R., Succop, P. Levin, L., Welge, J., dan Sivaganesan, S. (2005). Statistical Approaches to Analyze Censored Data with Multiple Detection Limits. *Disertasi Program Doctor of Philosophy University of Cincinnati.*