Pendugaan Parameter Distribusi Kasus Positif COVID-19 di Indonesia dengan Metode Bayes

Inka Rizky Sumantri*, Abdul Kudus

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*inkakarizky98@gmail.com, Abdul.kudus@unisba.ac.id

Abstract. Inference statistics is a statistical method that is useful for inferring population data from analyzes carried out by withdrawing the sample data. Parameter estimation is one part of statistical inference. The parameter estimation steps in inferencing statistics can be done using two methods, namely the classical method or the Bayes method. Bayes method uses the sample distribution and prior distribution to determine the posterior distribution. The data used is weekly extreme data for positive cases of Covid-19 in Indonesia. By using Kolmogorov Smirnov test, the data follows a Normal distribution. The conjugate prior of unknown $\boldsymbol{\mu}$ and known $\boldsymbol{\sigma^2 = \theta}$ is $\boldsymbol{\mu \sim N(\mu_0, \theta_0)}$ and the posterior distribution is $\boldsymbol{N(\mu', \theta')}$. The conjugate prior of known $\boldsymbol{\mu}$ and unknown $\boldsymbol{\sigma^2 = \theta}$ is Inverse Gamma $\boldsymbol{\theta \sim \mathrm{IG}(\alpha_0, \kappa_0)}$ the posterior distribution is $\boldsymbol{IG\ (\alpha', k')}$. The results of estimating parameters using the Bayes Markov Chain Monte Carlo method, the Gibbs Sampling algorithm shows that the average estimated value is 1.978.347 and the sigma value (σ) is 1.637.65.

Keywords: Bayes Method, Markov Chain Monte Carlo, Gibbs Sampling, Covid-19.

Abstrak. Statistika inferensia adalah metode statistika yang berguna untuk menyimpulkan data populasi dari analisis yang dilakukan dengan penarikan data sampelnya. Pendugaan parameter merupakan salah satu bagian dari inferensi statistika. Tahapan pendugaan parameter dalam statistika inferensia dapat dilakukan menggunakan dua cara, yaitu dengan metode klasik atau metode Bayes. Salah satu metode yang digunakan adalah metode Bayes yang menggunakan distribusi sampel dan distribusi prior untuk menentukan distribusi posteriornya. Data yang digunakan merupakan data ekstrim mingguan untuk kasus positif Covid-19 di Indonesia. Menggunakan uji kecocokan distribusi Kolmogorov Smirnov data kasus positif COVID-19 maksimum mingguan di Indoneisa mengikuti distribusi Normal. Prior konjugat dari $\boldsymbol{\mu}$ tidak diketahui dan $\boldsymbol{\sigma^2 = \theta}$ diketahui adalah $\boldsymbol{\mu \sim N(\mu_0, \theta_0)}$ dan distribusi posteriornya adalah $\boldsymbol{N \sim (\mu', \theta')}$. Sedangkan untuk $\boldsymbol{\mu}$ diketahui dan $\boldsymbol{\sigma^2 = \theta}$ tidak diketahui distribusi priornya adalah Inverse Gamma atau $\boldsymbol{\theta \sim \mathrm{IG}(\alpha_0, \kappa_0)}$ dan distribusi posteriornya adalah $\boldsymbol{IG \sim (\alpha', k')}$. Berikutnya hasil pendugaan parameter menggunakan metode Bayes Markov Chain Monte Carlo algoritma Gibbs Sampling didapat bahwa nilai pendugaan rata-ratanya sebesar 1.978,347 dan nilai sigma $(\boldsymbol{\sigma})$ sebesar 1.637,65.

Kata kunci: Metode Bayes, Markov Chain Monte Carlo, Gibbs Sampling, Covid-19.

1. Pendahuluan

Statistika adalah salah satu disiplin ilmu yang mempelajari tentang data, baik cara mendapatkan data, pengolahan data, analisis data, dan menyimpulkan data yang telah dianalisis (Sudjana, 2005). Statistika dibagi menjadi dua bagian besar yaitu, statistika deskriptif yang berhubungan dengan pengelompokan data, peringkasan data, dan penyajian dari data yang ada hingga menjadi informasi yang berguna dan mudah untuk dipahami dan statistika inferensia yang berguna untuk menyimpulkan data populasi dari analisis yang dilakukan dengan penarikan data sampelnya. Dalam pendugaan parameter harus diidentifikasi lebih dulu distribusi dari sebuah sampel acak dengan menggunakan beberapa metode.

Tahapan pendugaan parameter dalam statistika inferensia dapat dilakukan dengan dua cara, yaitu dengan metode klasik atau metode Bayes. Metode klasik mengandalkan proses inferensi pada data sample yang telah diambil dari populasi sementara metode Bayes turut serta memperhitungkan distribusi awal yang disebut prior pada analasis data sampel yang sudah diambil dari suatu populasi. Kesulitan yang dihadapi saat menggunakan metode Bayes adalah dalam perhitungan integral yang kompleks yang dilakukan untuk menentukan distribusi posterior. Distribusi posterior yang terbentuk kemudian akan diduga parameter-parameternya dengan menggunakan metode Rantai Markov Monte Carlo (MCMC) (Mukid dan Wulandari, 2012).

Covid-19 merupakan penyakit baru yang muncul dan mudah menyebar ke seluruh dunia. Penyebarannya cairan tubuh (*droplet*) seperti bersin. Di Indonesia, kasus pertama Covid-19 terdeteksi pada tanggal 2 Maret 2020 dengan 2 orang yang positif Covid-19 dan terus betambah setiap harinya.

Jika melihat data maksimum mingguan kasus positif COVID-19 di Indonesia, maka metode analisis Bayes dapat digunakan untuk menduga parameter dari suatu fungsi distribusi peluang. Langkah-langkah yang dilakukan adalah mencari fungsi distribusi yang cocok dengan data Covid-19, mencari distribusi prior dan distribusi posteriornya sehingga dapat menduga parameter distribusi pada kasus positif baru Covid-19 maksimum mingguan. Tujuan dalam penelitian ini adalah untuk menduga parameter dari data maksimum mingguan kasus positif COVID-19 di Indonesia.

2. Landasan Teori

Uji Kecocokan Distribusi

Uji Kolmogorov Smirnov sering disebut juga uji kecocokan non parametrik, karena pengujiannya tidak menggunakan fungsi distribusi tertentu. Uji ini digunakann untuk menguji selisih terbesar antara peluang pengamatan (empiris) dengan peluang teoritis (Fajarani dkk, 2018). Tahapan uji Kolmogorov Smirnov adalah sebagai berikut :

1. Hipotesis
   H_o: S(x) = F_0(x)
   H_1: S(x) ≠ F_0 (x)
2. Statistik uji
   $D = \sup_x | S(x) - F_0(x) |$
   dimana :
   S(x) = fungsi distribusi sampel (empiris)
   F_0(x) = fungsi distribusi yang dihipotesikan
3. Kriteria uji Menolak H_0 apabila nilai p-value < α atau nilai $D > D_\alpha$

Nilai Ekstrim

Metode Blok Maksima dapat digunakan untuk menentukan nilai-nilai ekstrim menurut periode tertentu misalnya menentukan nilai maksimum atau minimum dari suatu pengamatan perhari, perminggu, perbulan, dan pertahun (Bisono, 2011). Dalam pendekatan ini data harian kasus positif COVID-19 di Indonesia dapat mengambil nilai maksimum dari periode waktu tiap minggunya.

Dalam penelitian yang dilakukan oleh Fisher dan Tippett dengan pendekatan metode Blok Maksima menunjukkan kecenderungan mengikuti salah satu dari tiga sebaran dasar nilai ekstrim. Ketiga bentuk sebaran yang dimaksud adalah sebaran Gumbel, sebaran Frechet, dan sebaran Weibull (Bisono, 2011).

Gumbel: $H(x) = \exp\left\{-\exp\left[-\left(\frac{x-b}{a}\right)\right]\right\}; -\infty < x < \infty$ ...(2.1)

Frechet: $H(x) = \begin{cases} 0, \ x \leq a \\ \exp\left\{-\left(\frac{x-b}{a}\right)^{-a}\right\}, \ x > a \end{cases}$ ...(2.2)

Weibull: $H(x) = \begin{cases} \exp\left\{-\left[-\left(\frac{x-b}{a}\right)\right]^{a}\right\}, \ x < a \\ 1, \ x \geq a \end{cases}$ ...(2.3)

Ketiga distribusi ekstrim memiliki bentuk ujung sebaran yang berbeda-beda. Perbedaan ini memberikan gambaran yang berbeda untuk perilaku nilai ekstrim, maka akan sulit untuk menentukan secara tepat pola sebaran dari nilai ekstrim. Permasalahan ini bisa diselesaikan dengan menggabungkan ketiga tipe sebaran ke dalam sebaran nilai ekstrim (*generalized extreme value*/GEV) sebagai berikut:

$G(x) = exp\left(-\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-1/\xi}\right)$ ...(2.4)

Pendugaan Parameter

Pendugaan merupakan proses statistik yang digunakan untuk menghasilkan sebuah hasil dugaan (*estimate*) dari suatu parameter. Penduga statistik (rata-rata, persentase, variasi, dan lainnya) digunakan untuk mengestimasi sebuah parameter (Harinaldi dalam Rarasati, 2012). Sifat-sifat penduga adalah sebagai berikut:

1. Tidak Bias
   Penduga harus mendekati nilai sebenarnya dari parameter yang ada tersebut. Misalkan terdapat θ sebagai parameter. Jika $\hat{\theta}$ merupakan penduga tak bias dari parameter θ, maka nilai $E(\hat{\theta})$ harus sama dengan θ (Misbahussrur, 2009).
2. Efisien
   Penduga yang efisien adalah jika memiliki varians yang kecil.
3. Konsisten
   Suatu penduga dikatakan konsisten apabila jika ukuran sampel bertambah maka penduga akan mendekati parameternya.

Metode Bayes

Metode Bayes adalah suatu metode yang menggunakan data historis (data prior) untuk melakukan penelitian saat ini. Data historis (data prior) harus diketahui terlebih dahulu distribusinya sebelum merumuskan distribusi posteriornya. Selanjutnya distribusi prior dan distribusi posterior dapat membantu untuk menyelesaikan bagian yang sulit dari sebuah solusi (Rarasati, 2012).

Misalkan ruangan sampel S dipartisi menjadi kejadian yang *mutually exclusive* dan *exhaustive* $A_1, A_2, A_3, \ldots, A_K$ dengan $P(A_i) \neq 0$ untuk $i = 1,2,3, \ldots, K$. Misalkan terdapat kejadian B di dalam ruangan sampel S sehingga $P(B) > 0$ maka untuk sembarang kejadian B,

$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+\cdots+P(B|A_K)P(A_K)}$ ...(2.5)

Persamaan di atas merupakan persamaan Bayes.

Misalkan $X_1, X_2, X_3, \ldots, X_n$ merupakan sebuah sampel acak berukuran n dari distribusi yang mempunyai fungsi densitas peluang berbentuk $f(x|\theta)$. Langkah-langkah untuk menentukan estimasi Bayes bagi θ adalah (Heryanto, 2011) :

1. Penentuan fungsi densitas peluang gabungan dari sampel acak $X_1, X_2, X_3, \ldots, X_n$

$f(x_1, x_2, x_3, \dots, x_n|\theta) = f(x_1|\theta)f(x_2|\theta)f(x_3|\theta)\dots f(x_n|\theta)$ ...(2.6)

2. Penentuan fungsi densitas dari θ, yang besarnya dipilih dan disesuaikan dengan $f(x_1, x_2, x_3, \dots, x_n)$. Distribusi yang mempunyai fungsi densitas dari θ, dinotasikan $\lambda(\theta)$, dinamakan distribusi prior.
3. Penduga Bayes untuk θ ditentukan oleh:
   $T = E_{\theta|x}[\theta] = \int \theta f(\theta|x)\, d\theta$ ...(2.7)
   Distribusi posterior ditentukan dengan:
   $f(\theta|x) = \frac{f(x_1,x_2,x_3,\dots,x_n|\theta)\lambda(\theta)}{\int f(x_1,x_2,x_3,\dots,x_n|\theta)\lambda(\theta)d\theta}$ ...(2.8)

Fungsi Likelihood

Fungsi *likelihood* adalah fungsi densitas bersama dari *n* variabel acak X_1, X_2, $X_{3, \dots,}$ X_n dan dinyatakan dalam bentuk $f(x_1, x_2, x_3, \dots, x_n; \theta)$. Jika $x_1, x_2, x_3, \dots, x_n$ tetap, maka fungsi *likelihood* adalah fungsi dari parameter θ dan dinotasikan dengan $L(\theta)$. Jika $x_1, x_2, x_3, \dots, x_n$ menyatakan suatu sampel acak, maka:

$L(\theta) = f(x_1, x_2, x_3, \dots, x_n; \theta)$
$= f(x_1; \theta)f(x_2; \theta)f(x_3; \theta)\dots f(x_n; \theta)$
$= \prod_{i=1}^{n} f(x_i; \theta)$ ...(2.9)

Penduga *maximum likelihood* dari θ adalah suatu nilai $\hat{\theta}$ yang akan memaksimumkan fungsi *likelihood* $L(\theta)$. Untuk memudahkan proses komputasi, maka $L(\theta)$ dapat ditransformasi menggunakan logaritma natural *ln*, karena nilai θ yang memaksimumkan ln $L(\theta)$ sama dengan nilai θ yang memaksimumkan $L(\theta)$.

$ln\, L(\theta) = ln(\prod_{i=1}^{n} f(x_i; \theta)) = \sum_{i=1}^{n} lnf(x_i; \theta)$ ...(2.10)

Nilai θ yang memaksimumkan $ln\, L(\theta)$, diperoleh dengan mendifferensialkan $ln\, L(\theta)$ terhadap θ dan menyamakannya dengan 0, dan memastikan bahwa turunan keduanya kurang dari 0. Nilai θ yang memaksimumkan $ln\, L(\theta)$ disebut sebagai penduga *maximum likelihood* dari θ dan dinotasikan dengan $\hat{\theta}$.

Distribusi Prior

Distribusi prior adalah bentuk distribusi yang merupakan representasi objektif pada suatu parameter yang lebih rasional untuk dipercayai (Mulyaningsih, 2016). Dalam analisis Bayes, jika suatu populasi mengikuti distribusi tertentu dengan suatu parameter di dalamnya, maka dimungkinkan bahwa parameter itu sendiri juga mengikuti suatu distribusi probabilitas tertentu, yang disebut sebagai distribusi prior. Rahmadiah mengatakan distribusi prior dibagi menjadi dua, yaitu:

1. Berkaitan bentuk distribusi hasil identifikasi pola datanya diperoleh dari fungsi *likelihood*, yaitu:
   Distribusi prior konjugat yang mengacu pada penelitian prior pada suatu model terutama dalam pembentukan pola fungsi *likelihood*nya.
   Distibusi prior non konjugat. Pemberian prior pada model tidak mengindahkan pola pembentuk fungsi *likelihood*.
2. Berkaitan dengan informasi terdahulu terkait dengan penentuan masing-masing parameter pada pola distribusi priornya, yaitu:
   Distribusi prior informatif. Distribusi prior ini mengacu pada pemberian parameter dari distribusi prior yang telah dipilih, baik distribusi prior konjugat atau prior non-konjugat. Pemberian nilai parameter pada distribusi prior ini mempengaruhi bentuk distribusi posterior yang akan didapatkan pada informasi data yang diperoleh.
   Distribusi prior non-informatif. Distribusi prior ini tidak berdasarkan pada data yang ada atau distribusi prior yang tidak mengandung informasi mengenai parameter.

Distribusi Posterior

Distribusi posterior akan sebanding dengan perkalian *likelihood* dari data dengan distribusi prior (Yunita dkk, 2016). Fungsi kepadatan bersama dan marginal yang diperlukan dapat ditulis dalam bentuk distribusi prior dan fungsi *likelihood*.

$f(\theta|x_i) = f(\theta)L(\theta)$ ...(2.11)

Fungsi Distribusi Normal
Banyak metode dan teknik analisis menggunakan distribusi normal termasuk memodelkan nilai maksimum curah hujan (Mukid dan Wulandari, 2012). Bentuk kurva yang simetris menyerupai lonceng membuat distribusi ini sering digunakan saat melakukan analisis statistika. Bentuk kurva distribusi Normal dipengaruhi oleh dua parameter yaitu, simpangan baku (σ) dan rata-rata (μ). Nilai dari σ menentukan bentangan dari kurva sedangkan nilai dari μ menentukan pusat simetrisnya. Fungsi kepadatan peluang variabel random Normal X, dengan rata-rata μ dan varians σ^2, adalah sebagai berikut:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\ -\infty < x < \infty \qquad \text{... (2.12)}$$

Apabila suatu sampel acak berdistribusi Normal dengan menggunakan likelihoodnya maka distribusi prior konjugat dari parameter μ mengikuti distribusi Normal dan parameter σ^2 mengikuti distribusi *Inverse Gamma*.

Markov Chain Monte Carlo
Markov Chain Monte Carlo atau MCMC adalah metode simulasi dengan menggabungkan antara Monte Carlo dan sifat Markov Chain guna mengatasi integral posterior yang rumit. MCMC digunakan untuk mendapatkan data sampel yang didasari oleh sampling tertentu (Ntzoufras, 2009). Perhitungan MCMC ini untuk menghasilkan distribusi posterior dari suatu persamaan Bayes yang rumit. Integrasi Monte Carlo dapat digunakan untuk perkiraan distribusi posterior (atau marginal posterior) yang diperlukan untuk analisis Bayes. Terdapat beberapa algoritma yang digunakan dalam metode MCMC seperti *Metropolish Hasting* dan *Gibbs Sampling*.

Gibbs Sampling
Gibbs sampling merupakan algoritma yang sangat efisien, sehingga sering digunakan sebagai generator variabel acak pada analisis data yang menggunakan metode MCMC (Iriawan, 2011). *Gibbs sampling* adalah teknik terkenal untuk menghasilkan sampel dari distribusi multivariat rumit yang sering digunakan dalam prosedur Monte Carlo.

Langkah-langkah algoritma *Gibbs Sampling.* Misalkan $\boldsymbol{x} = (x_1, x_p, \dots, x_p)$
Langkah 1: Menentukan nilai awal $x^{(0)}$
Langkah 2: Untuk t = 1,2,..., T ulangi langkah-langkah berikut

1. Menentukan $\boldsymbol{x} = \boldsymbol{x^{(t-1)}}$
2. untuk j=1,...,p perbarui x_j dari $x_j \sim f(x_j|\boldsymbol{x}_{\backslash j})$ dimana $\boldsymbol{x}_{\backslash j} = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_p)^T$. Proses lengkapnya sebagai berikut:

$x_1{}^{(t)}$ dari $f(x_1|x_2{}^{(t-1)}, x_3{}^{(t-1)}, \dots, x_p{}^{(t-1)})$
$x_2{}^{(t)}$ dari $f(x_2|x_1{}^{(t)}, x_3{}^{(t-1)}, \dots, x_p{}^{(t-1)})$
$x_3{}^{(t)}$ dari $f(x_3|x_1{}^{(t)}, x_2{}^{(t)}, x_4{}^{(t-1)} \dots, x_p{}^{(t-1)})$
...
$x_j{}^{(t)}$ dari $f(x_j|x_1{}^{(t)}, x_2{}^{(t)}, \dots, x_{j-1}{}^{(t)}, x_{j+1}{}^{(t-1)}, \dots, x_p{}^{(t-1)})$
...
$x_p{}^{(t)}$ dari $f(x_p|x_1{}^{(t)}, x_2{}^{(t)}, \dots, x_{p-1}{}^{(t)})$

3. membentuk $x^{(t)}$ dan menyimpannya sebagai himpunan nilai-nilai yang dibangkitkan pada iterasi ke (t+1) dari algoritma.

Membangkitkan nilai-nilai dari $f(x_j|\boldsymbol{x}_{\backslash j}) = f(x_j|\, x_1{}^{(t)}, \dots, x_{j-1}{}^{(t)}, x_{j+1}{}^{(t-1)}, \dots, x_p{}^{(t-1)})$ cukup mudah karena merupakan distribusi univariat dimana semua variabel-variabel kecuali x_j dipertahankan tetap pada nilai-nilai yang diberikannya (Ntzoufra, 2009).

COVID-19

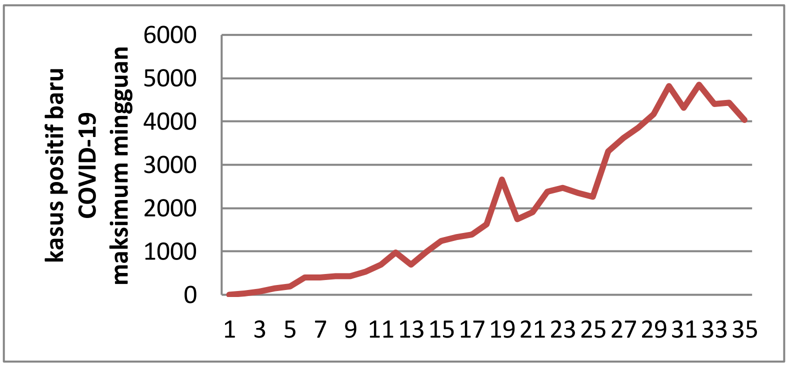
Coronavirus merupakan virus baru ditemukan pertama kali di Kota Wuhan Cina, pada Desember 2019 yang kemudian dinamakan *Severe Acute Respiratory Syndrome Coronavirus* 2 (SARS-CoV-2), dan menyebabkan penyakit *Coronavirus Deases*-2019 (COVID-19).

Sedangkan kasus pertama yang tercatat di Indonesia tanggal 2 Maret 2020 dengan kasus pasien positif sebanyak 2 orang. Penyebaran virus Corona sangat cepat sehingga dalam kurun waktu beberapa bulan sudah menyebar ke seluruh dunia. Seperti penyakit pada umumnya, COVID-19 juga memiliki gejala yang timbul seperti demam di atas 38 derajat. Namun ada pula yang tidak memiliki gejala tetapi masih bisa menularkan virus pada orang lain. Meski begitu, angka kematian COVID-19 masih lebih kecil dibandingkan dengan penyakit SARS yang muncul pada tahun 2003, yaitu kurang dari 5%.

3. Hasil Penelitian dan Pembahasan

Analisis Deskriptif

Pada grafik terlihat bahwa kasus positif baru COVID-19 terus meningkat. Hal ini dikarenakan bahwa pandemi masih belum bisa dikendalikan. Kasus pertama di Indonesia terjadi pada tanggal 2 Maret dengan 2 kasus positif baru COVID-19. Dapat dilihat kasus tertinggi terjadi pada minggu ke 32 dengan jumlah kasus mencapai 4.850, dan minggu setelahnya mengalami penurunan jumlah kasus. Pada minggu ke 33 terdapat 4.411 kasus dan terus menurun hingga minggu ke 35.



Gambar 1. Grafik Garis untuk Data Kasus Positif baru COVID-19 Maksimum Mingguan di Indonesia 2 Maret 2020 hingga 1 November 2020

Selama 35 minggu awal kasus positif baru COVID-19 maksimum mingguan di Indonesia memiliki rata-rata sebesar 1.977,5 dengan simpangan bakunya 1.587,7. Kasus positif baru COVID-19 maksimum mingguan terendah adalah sebanyak 2 kasus dan tertinggi sebanyak 4.850 kasus. Nilai kemencengan data (*skewness*) mencapai 0,45029 dan nilai kemenjuluran data (*kurtosis*) -1,2079.

Tabel 1. Statistika Deskriptif Kasus Positif Baru COVID-19 Maksimum Mingguan di Indonesia

Statistik	Nilai
Sample Size	35
Range	4.848
Mean	1.977,5
Std. Deviation	1.587,7
Coef. of Variation	0,80287
Std. Error	268,37
Skewness	0,45029
Excess Kurtosis	-1,2079

Sebelum menduga parameter akan dilakukan indentifikasi jenis distribusi terlebih dahulu. Menggunakan software Easyfit didapat kesimpulan dari uji Kolmogorov-Smirnov untuk setiap pengujian distribusi adalah sebagai berikut.

Tabel 2. Kesimpulan Uji Kolmogorov-Smirnov

Distribusi	Nilai Statistik Hitung	Kriteria Uji	Keterangan
Gumbel	0,13116	< 2,24	Terima H_0
Weibull	0,18327	< 2,24	Terima H_0
Frechet	0,23874	> 2,24	Tolak H_0
Normal	0,12946	< 2,24	Terima H_0

Didapat kesimpulan bahwa data kasus positif baru COVID-19 maksimum mingguan di Indonesia paling cocok mengikuti distribusi Normal karena memiliki nilai statistik hitung yang lebih kecil dibanding nilai kritisnya dan lebih kecil dibanding dengan distribusi Gumbel dan Weibull.

Pendugaan Parameter Distribusi

Diketahui bahwa sebaran data kasus positif baru COVID-19 maksimum mingguan di Indonesia mengikuti distribusi Normal dengan parameter rata-rata (μ) dan ragam (σ^2). Mengikuti prior konjugatnya maka apabila μ tidak diketahui dan $\sigma^2 = \theta$ diketahui maka prior konjugatnya berdistribusi Normal, yakni $\mu \sim N(\mu_0, \theta_0)$, dan apabila μ diketahui dan $\sigma^2 = \theta$ tidak diketahui maka prior konjugatnya berdistribusi Inverse Gamma atau $\theta \sim \mathrm{IG}(\alpha_0, \kappa_0)$. Masing-masing memiliki fungsi densitas distribusi prior sebagai berikut:

$$f(\mu; \mu_0, \theta_0) = \frac{1}{\sqrt{2\pi\theta_0}} e^{-\frac{1}{2\theta_0}(x_i - \mu_0)^2} \qquad ...(3.32)$$

$$p(\theta; \alpha_0, k_0) = \frac{{k_0}^{\alpha_0}}{\Gamma(\alpha_0)} \theta^{-(\alpha_0+1)} e^{\left(-\frac{k_0}{\theta}\right)} \qquad ...(3.33)$$

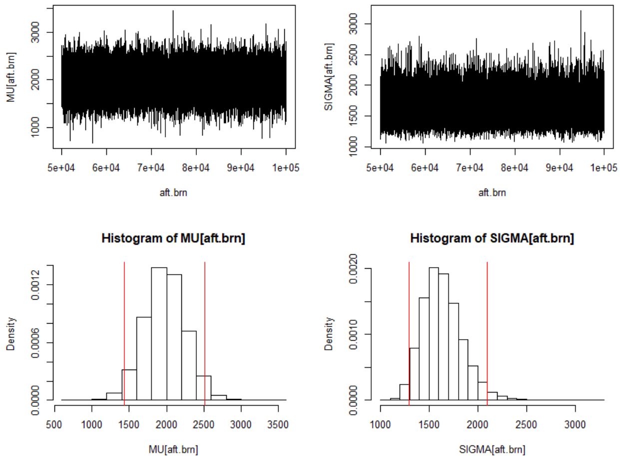
Distribusi posterior adalah penggabungan antara distribusi prior dengan distribusi sampel. Distribusi posterior sebanding dengan perkalian dari distribusi likelihood data dan distribusi prior.

PRIOR × LIKELIHOOD ∝ POSTERIOR

Diketahui bahwa distribusi data adalah normal $N\ (\mu, \sigma^2 = \theta)$ maka distribusi rata-rata dari sampel acak berukuran *n* adalah normal $N\ \left(\mu, \frac{\theta}{n}\right)$. Distribusi prior konjugat untuk μ tidak diketahui dan $\sigma^2 = \theta$ diketahui adalah Normal $N(\mu_0, \theta_0)$. Maka hasil kali dari distribusi prior dan *likelihood*nya akan mengikuti distribusi Normal $N\ \left(\mu' = \theta' \left(\frac{\mu_0\theta + n\bar{x}\theta_0}{\theta_0\theta}\right),\ \theta' = \frac{\theta_0\theta}{\theta + n\theta_0}\right)$. Sedangkan distribusi posterior untuk parameter $\sigma^2 = \theta$ dengan distribusi priornya adalah *Inverse Gamma* adalah $IG \sim \left(\alpha' = \alpha_0 + {}^{n}/_{2}, k' = k_0 + \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{2}\right)$.

Gibbs Sampling

Proses membangkitkan sampel posterior menggunakan *Software R* dengan algoritma *Gibbs Sampling* akan terus dilakukan secara iteratif sampai konvergen. Sesuai dengan hasil statistik deskriptif dalam tabel 1, maka dalam penelitian ini digunakan nilai awal bagi $\mu_0 = 2000$ dan nilai awal bagi $\theta_0 = 2500000$ dan nilai $\alpha_0 = \frac{1}{6}, k_0 = \frac{1}{2}$ dengan data sebanyak 35 pengamatan. Kemudian Pemantauan hasil dari konvergensi iterasi dapat dilihat di Gambar 2.



Gambar 2. *Trace plot* dan Histogram Pendugaan Parameter Rata-rata (μ) dan Sigma (σ)

Trace plot menunjukkan plot banyaknya iterasi yang dijalankan dengan sampel yang dibangkitkan. Pada Gambar 2 bagian atas terlihat bahwa *trace plot* bagi penduga kedua parameter sudah tidak membentuk suatu pola atau *trend*. Hal ini menandakan bahwa *burn-in period* sudah selesai artinya sampel yang dibangkitkan sudah berada di daerah distribusi target. Maka dapat dikatakan bahwa algoritma mencapai konvergensi. *Trace plot* tersebut juga menunjukkan hasil dari penduga rata-rata (μ) dan sigma (σ) dapat dipercaya sebagai nilai hasil yang akurat dari distribusi posterior.

Dari hasil pendugaan didapat nilai rata-rata (μ) dari hasil pendugaan di atas adalah sebesar 1.978,347 dengan Bayes interval sebesar 5% terletak di antara 1.435,386 sampai 2.516,441. Sedangkan nilai sigma (σ) dari hasil pendugaan di atas adalah sebesar 1.637,65 dengan Bayes interval sebesar 5% terletak di antara 1.292,830 sampai 2.097,546.

4. Kesimpulan

Menggunakan metode Bayes Markov Chain Monte Carlo diperoleh nilai dugaan untuk parameter rata-rata distribusi Normal adalah 1.978,347 dengan simpangan baku sebesar 1.637,65. Hal ini menunjukkan bahwa rata-rata kasus positif baru COVID-19 maksimum mingguan di Indonesia sekitar 1.978,347 kasus atau 1.978 orang dinyatakan positif. Selanjutnya inferensi mengenai kasus baru positif maksimum mingguan COVID-19 di Indonesia dapat dilakukan berdasarkan fungsi distribusi Normal (μ = 1978,347, σ = 1637,65).

5. Saran

Penelitian ini masih dibutuhkan pengembangan keilmuan oleh karenanya disarankan untuk penelitian selanjutnya menggunakan metode Bayes Markov Chain Monte Carlo algoritma *Gibbs Sampling* dengan menggunakan distribusi prior lain. Selain itu dapat juga menduga parameter distribusi lain menggunakan metode Bayes Markov Chain Monte Carlo.

Daftar Pustaka

[1] Anifa., Mukid, M.A., Rusgiyono, Agus. (2012). Simulasi Stokastik Menggunakan Algoritma Gibbs Sampling. *Jurnal Gaussian*. **1**(1), 21-30.

[2] Bisono, Indriati,. (2011). Mengenal Data Ekstrim dan Distribusinya. *Jurnal Teknik Industri*. **13**(2), 81-86.

[3] Fajarani, G.I., Purnamasari, Ika., Wahyuningsih, Sri. (2018). Prediksi Data Curah Hujan Dengan Menggunakan Statistika Non Parametrik. *Jurnal Eksponensial*. **9**(2), 177-186.

[4] Heryanto, N. (2011). *Statistika Matematis Lanjutan*. Bandung: CV Pustaka Setia.

[5] Iriawan, N. (2011). *Pemodelan Mixture of Mixture dalam Pemilihan Portofolio*, Prosiding Seminar Nasional Statistika Universitas Diponegoro, Semarang.

[6] Misbahussurur, Ahmad. (2009). *Estimasi Parameter Distribusi Gamma Dengan Metode Maksimum Likelihood*. Skripsi tidak diublikasikan. Malang: Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Ibrahim Malang.

[7] Mukid, M.A., Wilandari, Yuciana. (2012). Identifikasi Pola Distribusi Curah Hujan Maksimum Dan Pendugaan Parameternya Menggunakan Metode Bayesian Markov Chain Monte Carlo. *Media Statistika*. **5**(2), 63-74.

[8] Mulyaningsih, M.D. (2016). *Estimasi Parameter Distribusi Binomial Negatif Dengan Pendekatan Bayesian Menggunakan Monte Carlo Markov Chain Berdasarkan Algoritma Mteropolish Hasting*. Skripsi tidak dipublikasi. Surabaya: Program Studi Statistika Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga.

[9] Ntzoufras, I. (2009). *Bayesian Modelling using WinBUGS*. USA: Wiley.

[10] Rarasati, I.P. (2012). *Estimasi Parameter Distribusi Gamma Dengan Metode Bayes*. Skripsi tidak diublikasikan. Malang: Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Ibrahim Malang.

[11] Sudjana. (2013). *Metoda Statistika*. Edisi ke-13. Bandung: PT. Tarsito Bandung.

[12] Yunita, Rini dan Subanar, Abdurrakhman. (2016). Estimasi Bayesian Pada Model Persamaan Struktural Dengan Variabel Kategorik Terurut. *Jurnal IPTEKS Terapan*. **10**(2), 86-94.