

PENYELESAIAN NUMERIK UNTUK MENENTUKAN
NILAI OPTIMAL PADA *AMERICAN OPTION*
DENGAN METODE BEDA HINGGA *FULLY*
IMPLISIT DAN CRANK-NICOLSON

Lukman Hanafi, Endah Rohmati M.P., G. M. Puspita
Jurusan Matematika FMIPA-ITS Surabaya
email: lukman@matematika.its.ac.id

Abstrak

Option adalah sebuah kontrak keuangan yang memberikan hak kepada pemiliknya untuk membeli atau menjual sejumlah aset dasar dalam jangka waktu tertentu sesuai dengan harga yang disepakati pada saat penandatanganan kontrak *option*. *American option* sifatnya dapat di*exercise* sewaktu-waktu mulai dari penandatanganan kontrak sampai jatuh tempo. Melakukan *exercise* pada saat optimal adalah strategi yang dilakukan *holder* untuk mendapatkan keuntungan yang maksimal, namun waktu yang tepat untuk mendapatkan hasil yang optimal tersebut belum diketahui (*free boundary-value problem*). Penentuan harga optimal *American option* berdasarkan *free boundary-value problem* dilakukan pendiskritan persamaan Black-Scholes menggunakan metode Beda Hingga *Fully* Implisit dan Crank-Nicolson. Selanjutnya, kedua penyelesaian tersebut dibandingkan didapat bahwa Metode Beda Hingga Crank-Nicolson lebih baik berdasarkan kestabilan, konvergensi, dan akurasi.

Katakunci: *American option, persamaan Black-Scholes, metode beda hingga*

1. Pendahuluan

Option adalah sebuah kontrak yang memberikan hak (bukan suatu kewajiban) kepada pemiliknya (*holder*) untuk membeli (*call*) atau menjual (*put*) sejumlah aset dasar (*underlying asset*) dalam jangka waktu tertentu (*maturity date/expiration date*) sesuai dengan harga yang disepakati pada saat penandatanganan kontrak *option* (*strike price*) [4].

Dari sifat *early exercise*, *American option* memberikan *holder* hak yang lebih besar dibandingkan dengan *European option*, sehingga dengan *early exercise* tersebut *American Option* menjadi lebih bernilai (*valuable*) daripada *European option* [3], [6]. Atau dapat dinotasikan $VPamP = VPeurP$ artinya nilai dari *American option* selalu lebih besar atau sama dengan *European option* [9]. Melakukan *exercise* pada saat optimal adalah strategi yang dilakukan *holder* untuk mendapatkan keuntungan yang maksimal atau disebut juga strategi optimal *exercise*. Tentu saja waktu yang tepat untuk mendapatkan hasil yang optimal tersebut belum diketahui karena pada kenyataannya waktu yang optimal untuk meng*exercise option* berdasarkan harga saham yang optimal [8], [12]. Strategi optimal *exercise* ini termasuk dalam masalah dari *free boundary-value problem* [11].

Pada paper ini dikaji mengenai penentuan nilai optimal *American option* berdasarkan harga saham yang optimal yaitu saat harga saham yang dipilih mencapai nilai minimum untuk *put option* dan harga saham maksimum untuk *call option* dengan Metode Beda Hingga Fully Implisit dan Crank-Nicolson.

2. Persamaan Black-Scholes

Persamaan untuk *American option* tanpa pembagian dividen diberikan dalam persamaan Black-Scholes berikut [4], [7]:

$$\frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} + rS \frac{\partial f}{\partial S} - rf = 0 \quad (1)$$

Fungsi *payoff* untuk *American Option* diberikan:

$$\begin{cases} f(t, S) > \max(X - S, 0), \forall S > S_f(t) \\ f(t, S) = \max(X - S, 0), \forall S \leq S_f(t) \end{cases} \quad (2)$$

f : Nilai option (*option value*)

σ : Volatilitas

S : Harga saham saat waktu t

r : Suku bunga bebas resiko

X : *Strike price*

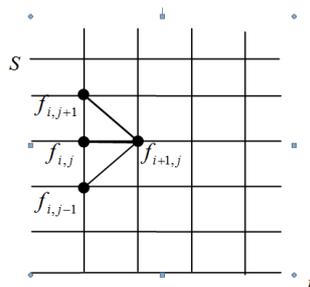
Persamaan Black-Scholes dengan waktu *exercise option* (T), dapat dibagi ke dalam N grid dengan panjang interval $\Delta t = \frac{T}{N}$. Total waktu (T) jika dilihat ke dalam grid yaitu $N + 1$ sehingga menjadi $0, \Delta t, 2\Delta t, \dots, T$

Misal S_{\max} adalah harga saham yang tertinggi saat di*exercise*, sehingga berdasarkan fungsi *payoff* untuk *put option* menjadi tidak bernilai. Hal tersebut didefinisikan $\Delta S = \frac{S_{\max}}{M}$ dan dimasukkan ke dalam grid yaitu $M + 1$ sehingga menjadi: $0, \Delta S, 2\Delta S, \dots, S_{\max}$

Jumlah titik-titik pada grid dari harga saham (S) dan waktu (t) adalah $(M + 1)(N + 1)$ titik. Koordinat titik (i, j) pada grid adalah titik yang menghubungkan antara waktu $i\Delta t$ dengan harga saham $j\Delta S$, sehingga untuk titik (i, j) dapat digunakan variabel $f_{i,j}$.

3. Pendiskritan Dengan Metode *Fully* Implisit

Metode Bada Hingga *Fully* Implisit adalah metode yang menggunakan pendekatan satu node penuh. Diskritisasi ke dalam bentuk *Fully* Implisit dilakukan dengan pendekatan beda maju. Pendekatan beda maju di titik $(i + 1, j)$ dan titik (i, j) sebagai titik interior, maka untuk Metode *Fully* Implisit dengan pendekatan maju diperoleh Gambar 1. Variabel-variabel



Gambar 1: Pembagian Grid pada Metode Bada Hingga Fully Implisit Bada Maju

yang akan digunakan dinotasikan sebagai

$$f(t, S) = f(i\Delta t, j\Delta S) = f_{i,j}$$

Dengan mengambil $S = j\Delta S$ dan $t = i\Delta t$ untuk setiap $j = 1, 2, \dots, M-1$ dan $i = 0, 1, \dots, N-1$ maka persamaan (1) akan didapat [1], [10]:

$$\begin{aligned} & \frac{f_{i+1,j} - f_{i,j}}{\Delta t} + r j \Delta S \left[\frac{f_{i,j+1} - f_{i,j-1}}{2\Delta S} \right] \\ & + \frac{1}{2} \sigma^2 j^2 (\Delta S)^2 \left[\frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{(\Delta S)^2} \right] - r f_{i,j} = 0 \end{aligned}$$

Persamaan ini disusun kembali menjadi :

$$\begin{aligned} f_{i+1,j} = & \left[\frac{1}{2} r j \Delta t - \frac{1}{2} \zeta \Delta t \right] f_{i,j-1} + [1 + \zeta \Delta t + r \Delta t] f_{i,j} \\ & + \left[-\frac{1}{2} r j \Delta t - \frac{1}{2} \zeta \Delta t \right] f_{i,j+1} \end{aligned}$$

dengan

$$\zeta = \sigma^2 j^2.$$

Persamaan ini ditulis

$$f_{i+1,j} = a_j f_{i,j-1} + b_j f_{i,j} + c_j f_{i,j+1} \quad (3)$$

dengan:

$$\begin{aligned} a_j &= \frac{1}{2} r j \Delta t - \frac{1}{2} \sigma^2 j^2 \Delta t \\ b_j &= 1 + \sigma^2 j^2 \Delta t + r \Delta t \\ c_j &= -\frac{1}{2} r j \Delta t - \frac{1}{2} \sigma^2 j^2 \Delta t \end{aligned}$$

Untuk setiap j merupakan grid yang membagi sisi S dengan interval $[0, M]$ yaitu sebanyak $M + 1$ dimensi vektor, demikian pula dengan i yang merupakan bagian dari grid yang membagi sisi t dengan interval $[0, N]$ yaitu sebanyak $N + 1$ dimensi vektor sehingga persamaan (3) jika ditulis dalam

perkalian matriks :

$$\begin{aligned}
 & \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_j & b_j & c_j & \cdots & 0 \\ 0 & a_j & b_j & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{N_{\min}, M_{\min}-1} & \cdots & f_{N_{\max}, M_{\min}} \\ f_{N_{\min}, M_{\min}} & & f_{N_{\max}, M_{\min}} \\ \vdots & & \\ f_{N_{\min}, M_{\max}-2} & \vdots & \vdots \\ f_{N_{\min}, M_{\max}-1} & \cdots & f_{N_{\max}, M_{\max}} \end{bmatrix} \\
 = & \begin{bmatrix} f_{N_{\min}+1, M_{\min}-1} & f_{N_{\min}+2, M_{\min}-1} & \cdots & BC \\ f_{N_{\min}+1, M_{\min}} & f_{N_{\min}+2, M_{\min}} & & BC \\ \vdots & & & \\ f_{N_{\min}+1, M_{\max}-2} & \vdots & & \vdots \\ f_{N_{\min}+1, M_{\max}-1} & f_{N_{\min}+2, M_{\max}} & \cdots & f_{N_{\max}, M_{\max}} \end{bmatrix} \quad (4)
 \end{aligned}$$

Syarat batas atau *Boundary Condition (BC)* diberikan persamaan (2) adalah :

$$f_{N+1,j} = \max[X - j\Delta S, 0], \text{ untuk setiap } j = 0, 1, 2, \dots, M$$

Nilai dari option saat harga saham 0 adalah X , sehingga:

$$f_{i,0} = X, \text{ untuk setiap } i = 0, 1, 2, \dots, M$$

Untuk $S = S_{\max}$ nilai option bernilai 0, sehingga;

$$f_{i,M+1} = 0, \text{ untuk setiap } j = 0, 1, 2, \dots, M$$

Persamaan (4) diatas dapat ditulis ke dalam persamaan kompak:

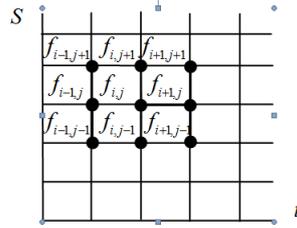
$$Af_{N,M} = f_{N+1,M}$$

Sehingga untuk mendapat nilai $f_{N,M}$ maka dapat dicari dengan mendapatkan nilai invers dari A dikalikan dengan $f_{N+1,M}$. Dengan demikian

$$f_{N,M} = A^{-1}f_{N+1,M}$$

4. Pendiskritan Metode Bada Hingga Crank-Nicolson

Metode Crank-Nicolson adalah rata-rata dari metode Eksplisit dan metode Implisit [2]. Pendekatan yang dilakukan adalah pendekatan beda maju dan pendekatan beda mundur. Titik-titik yang dipakai pada Metode Crank-Nicolson terlihat pada Gambar 2. Metode Crank-Nicolson dengan



Gambar 2: Pembagian Grid pada Metode Bada Hingga Crank-Nicolson

pendekatan beda maju diperoleh melalui rata-rata dari metode Implisit dan metode Eksplisit. Persamaan Black-Scholes pada persamaan (1) jika didiskritkan Dengan Metode Crank-Nicolson diperoleh [5] :

$$\begin{aligned} & \frac{(f_{i+1,j} - f_{i,j})}{\Delta t} + r S \left[\frac{1}{2} \frac{f_{i,j+1} - f_{i,j-1}}{2\Delta S} + \frac{1}{2} \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\Delta S} \right] \\ & + \frac{1}{2} \sigma^2 S^2 \left[\frac{1}{2} \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{(\Delta S)^2} + \frac{1}{2} \frac{f_{i+1,j+1} - 2f_{i+1,j} + f_{i+1,j-1}}{(\Delta S)^2} \right] \\ & - r \frac{1}{2} (f_{i+1,j} + f_{i,j}) = 0 \end{aligned}$$

Karena $S = j\Delta S$ dan $t = i\Delta t$ untuk setiap $j = 0, 1, 2, \dots, M$ dan $i = 0, 1, 2, \dots, N$ maka didapat:

$$f_{i+1,j} - f_{i,j} = a_j^* f_{i,j-1} + b_j^* f_{i,j} + c_j^* f_{i,j+1} + a_j^* f_{i+1,j-1} + b_j^* f_{i+1,j} + c_j^* f_{i+1,j+1}$$

dengan

$$\begin{aligned} a_j^* &= \frac{\Delta t}{4} (\sigma^2 j^2 - r j) \\ b_j^* &= -\frac{\Delta t}{2} (\sigma^2 j^2 + r) \\ c_j^* &= \frac{\Delta t}{4} (\sigma^2 j^2 + r j) \end{aligned}$$

Persamaan diubah kedalam suku-suku yang bersesuaian, didapat:

$$a_j^* f_{i+1,j-1} + (b_j^* + 1) f_{i+1,j} + c_j^* f_{i+1,j+1} = -a_j^* f_{i,j-1} - (b_j^* - 1) f_{i,j} - c_j^* f_{i,j+1} \quad (5)$$

Pendekatan dengan metode Crank-Nicolson :

$$a_j^* f_{i+1,j-1} + (b_j^* + 1) f_{i+1,j} + c_j^* f_{i+1,j+1} = g_{i,j}$$

dengan

$$g_{i,j} = -a_j^* f_{i,j-1} - (b_j^* - 1) f_{i,j} - c_j^* f_{i,j+1}$$

Untuk setiap j merupakan grid yang membagi sisi S dengan interval $[1, M + 1]$ yaitu sebanyak $M + 1$ dimensi vektor, begitu juga dengan i yang merupakan bagian dari grid yang membagi sisi t dengan interval $[1, N + 1]$ yaitu sebanyak $N + 1$ dimensi vektor sehingga persamaan (5) dapat ditulis dalam bentuk perkalian matriks:

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_j^* & b_j^* + 1 & c_j^* & \cdots & 0 \\ 0 & a_j & b_j^* + 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{N_{\min}, M_{\min}} & \cdots & f_{N_{\max}, M_{\min}} \\ f_{N_{\min}, M_{\min}+1} & & f_{N_{\max}, M_{\min}+1} \\ \vdots & & \vdots \\ f_{N_{\min}, M_{\max}} & & \vdots \\ f_{N_{\min}, M_{\max}} & \cdots & f_{N_{\max}, M_{\max}} \end{bmatrix} =$$

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -a_j^* & 1 - b_j^* & -c_j^* & \cdots & 0 \\ 0 & -a_j & 1 - b_j^* & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{N_{\min}-1, M_{\min}} & \cdots & BC \\ f_{N_{\min}-1, M_{\min}+1} & & BC \\ \vdots & & \vdots \\ f_{N_{\min}-1, M_{\max}-1} & & \vdots \\ f_{N_{\min}-1, M_{\max}} & \cdots & BC \end{bmatrix} \quad (6)$$

Persamaan (6) dapat ditulis bentuk ke dalam persamaan kompak :

$$Cf_{N+1,M} = Bf_{N,M}$$

Sehingga untuk mencari maka dapat dicari dengan mendapatkan nilai invers dari B dikalikan $Cf_{N+1,M}$. Dengan demikian :

$$f_{N,M} = B^{-1}Cf_{N+1,M}$$

Syarat batas atau *Boundary Condition (BC)* diberikan persamaan (2) adalah :

$$f_{N+1,j} = \max[X - j\Delta S, 0], \text{ untuk setiap } j = 0, 1, 2, \dots, M.$$

Nilai dari option saat harga saham 0 adalah X , sehingga:

$$f_{i,0} = X, \quad i = 0, 1, 2, \dots, M.$$

Untuk $S = S_{\max}$ nilai option bernilai 0, sehingga;

$$f_{i,M+1} = 0, \quad i = 0, 1, 2, \dots, M.$$

5. Tahap Perbandingan

Nilai optimal yang didapat dengan Metode *Fully* Implisit dan Crank-Nicolson untuk pembagian grid $N = M = 100, 200, \dots, 1900, 2000$ akan dibandingkan untuk mendapatkan metode mana yang lebih baik.

a. Akurasi dan Konvergensi

Nilai yang dipilih $S = 50$, $X = 55$, $S_{\max} = 100$, $T = \frac{5}{12}$, $\sigma = 0.3$, $r = 10\% = 0.1$, $\alpha = \frac{\Delta t}{(\Delta S)^2}$ dan nilai market untuk *American put option* didapat 7.0284

Nilai optimal dari Metode *Fully* Implisit dan Crank-Nicolson dapat dilihat pada Tabel 1.

Terlihat bahwa semakin banyak grid yang dipakai maka semakin kecil error yang terjadi terhadap nilai market untuk *American put option* sehingga semakin akurat nilai tersebut. Prosentase error dari metode *Fully* Implisit dengan error rata-rata 6.0188 adalah

$$\varepsilon_1^* = \frac{|7.0284 - 6.0188|}{7.0284} \times 100\% = 14.364578\%$$

Sehingga tingkat akurasi nilai optimal *option* adalah 85.635422%

Prosentase error dari metode Crank-Nicolson dengan error rata-rata 6.0226

$$\varepsilon_2^* = \frac{|7.0284 - 6.0226|}{7.0284} \times 100\% = 14.31051164\%$$

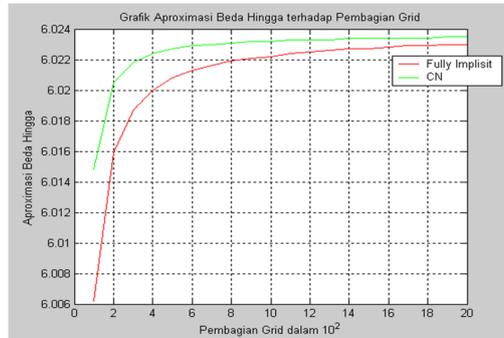
Tabel 1: Nilai optimal dari Metode *Fully* Implisit dan Crank-Nicolson

Grid	<i>Fully</i> Implisit		Crank-Nicolson		α
	Aprok simasi	Error	Aprok simasi	Error	
100	6.0062	1.0222	6.0148	1.0136	0.0042
200	6.0159	1.0125	6.0205	1.0079	0.0083
300	6.0187	1.0097	6.0218	1.0066	0.0125
400	6.0200	1.0084	6.0224	1.0061	0.0167
500	6.0208	1.0076	6.0227	1.0057	0.0208
600	6.0213	1.0071	6.0229	1.0056	0.0250
700	6.0216	1.0068	6.0230	1.0054	0.0292
800	6.0219	1.0065	6.0231	1.0053	0.0333
900	6.0221	1.0063	6.0232	1.0052	0.0375
1000	6.0222	1.0062	6.0232	1.0052	0.0417
1100	6.0224	1.0060	6.0233	1.0052	0.0458
1200	6.0225	1.0059	6.0233	1.0051	0.0500
1300	6.0226	1.0058	6.0233	1.0050	0.0542
1400	6.0227	1.0057	6.0234	1.0050	0.0583
1500	6.0227	1.0056	6.0234	1.0050	0.0604
1600	6.0228	1.0056	6.0234	1.0050	0.0625
1700	6.0229	1.0055	6.0234	1.0050	0.0708
1800	6.0229	1.0055	6.0234	1.0050	0.0750
1900	6.023	1.0054	6.0235	1.0049	0.0768
2000	6.023	1.0054	6.0235	1.0049	0.0773

Sehingga tingkat akurasi nilai optimal *option* adalah 85.6894883% Pada Gambar 3 perubahan nilai aproksimasi semakin konvergen menuju nilai Black-Scholes. Namun terdapat perbedaan pada Metode *Fully* Implisit dan Crank-Nicolson. Metode *Fully* Implisit konvergen pada titik 6.023 dan Metode Crank-Nicolson konvergen pada titik 6.0235. Jadi dapat disimpulkan bahwa dari perbandingan Akurasi dan Konvergensi Metode Crank-Nicolson lebih baik dibanding Metode *Fully* Implisit dalam penyelesaian persamaan Black-Scholes

b. Stabilitas

Untuk setiap $\alpha = \frac{\Delta t}{(\Delta S)^2}$ semakin besar ΔS maka semakin kecil nilai α , dan juga berlaku sebaliknya semakin kecil ΔS maka semakin besar nilai α untuk $0 < \alpha \leq 1$. Terlihat pada Gambar 4 bahwa perubahan nilai α terhadap nilai error dimana grafik *Fully* Implisit terletak di atas grafik Crank-Nicolson, sehingga dapat disimpulkan bahwa nilai Aproksimasi Metode Beda Hingga *Fully* Implisit dan Crank-Nicolson stabil. Metode Crank-Nicolson memiliki nilai α yang lebih kecil dibandingkan dengan Metode *Fully* Implisit pada grid yang sama. Sehingga dapat disimpulkan bahwa Metode Crank-Nicolson memiliki kestabilan yang lebih baik dibanding dengan Metode *Fully* Implisit.

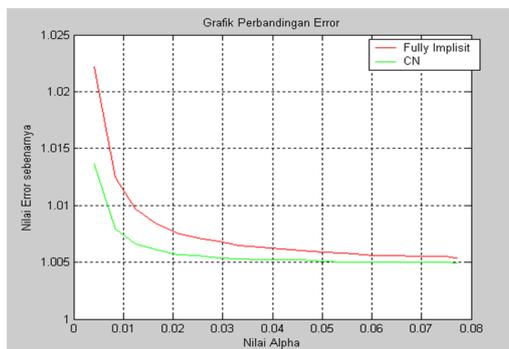


Gambar 3: Nilai Apoximasi Metode Bada Hingga

6. Kesimpulan

Penentuan nilai optimal *American put option* dilakukan melalui pendiskritan pada persamaan Black-Scholes dengan Metode *Fully* Implisit dan Metode Crank-Nicolson didapat persamaan diskrit: a. Metode *Fully* Implisit

$$f_{i+1,j} = a_j f_{i,j-1} + b_j f_{i,j} + c_j f_{i,j+1}$$



Gambar 4: Grafik Nilai Alpha Terhadap Error Sebenarnya

dengan :

$$\begin{aligned} a_j &= \frac{1}{2}r j \Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t \\ b_j &= 1 + \sigma^2 j^2 \Delta t + r \Delta t \\ c_j &= -\frac{1}{2}r j \Delta t - \frac{1}{2}\sigma^2 j^2 \Delta t \end{aligned}$$

b. Metode Crank-Nicolson

$$a_j^* f_{i+1,j-1} + (b_j^* + 1) f_{i+1,j} + c_j^* f_{i+1,j+1} = -a_j^* f_{i,j-1} - (b_j^* - 1) f_{i,j} - c_j^* f_{i,j+1}$$

dengan :

$$\begin{aligned} a_j^* &= \frac{\Delta t}{4} (\sigma^2 j^2 - rj) \\ b_j^* &= -\frac{\Delta t}{2} (\sigma^2 j^2 + r) \\ c_j^* &= \frac{\Delta t}{4} (\sigma^2 j^2 + rj) \end{aligned}$$

Untuk harga saham saat itu $S = 50$ harga saham $\max(S_{\max}) = 100$, strike price (X) = 55, volat (σ) = 0.3, suku bunga (r)=10% atau 0.1, pembagian grid untuk N diambil sama dengan M yaitu 100,200,...,1900,2000 didapat nilai optimal untuk Metode *Fully* Implisit adalah 6.023 dan untuk metode Crank-Nicolson adalah 6.0235.

Nilai optimal tersebut menunjukkan bahwa harga saham yang tepat untuk melakukan *exercise* option adalah saat harga saham berubah menjadi $(S + \frac{\Delta S}{M}) + 1$ untuk pembagian $N = M$ genap dan $S + \frac{\Delta S}{M}$ untuk pembagian $N = M$ ganjil.

Metode Crank-Nicolson lebih baik dibandingkan Metode *Fully* Implisit, karena dengan pembagian grid yang sama Metode Crank-Nicolson memiliki

nilai aproksimasi yang lebih mendekati nilai sebenarnya dibanding Metode *Fully* Implisit.

Pustaka

- [1] Basaruddin, T., *Metode Beda Hingga untuk Persamaan Differensial*, Jakarta: Elex Media Komputindo, 1994.
- [2] Brandimarte, Paolo, *Numerical Methods in Finance*, A Wiley-Interscience Publication, USA, 2002.
- [3] Dewinne, J., Howison, S., Willmot. P, *The Mathematics of Financial Derivatives*, Cambridge University Press, Australia, 1995.
- [4] Hull, John C., *Option Future and Other Derivatives*, Prentice Hall, New Jersey, 2002.
- [5] Kerman, J, *Numerical Methods for Option Pricing: Binomial and Finite-Difference Aproximaion*, New York University, USA, 2002.
- [6] Kusumaningtyas, K, *Estimasi Harga American Put Option Tanpa Dividen Menggunakan Metode Binomial*, Skripsi ITS, Surabaya, 2007.
- [7] Salim, Lani, *Derivatif : Option & Warrant* , Elex Media Komputindo, Jakarta, 2003
- [8] Sembel, Roy & Fardiansyah, Tedy, *Sekuritas Derivatif : madu atau racun?*, Salemba Empat, Jakarta, 2002.
- [9] Seydel, R, *Tools of Computation Finance*, Springer, Jerman, 2002.
- [10] Smith, G.D., *Numerical Solution of Partial Differential Equation: Finite Difference Methods*, Oxford University Press, New York, USA, 1985.
- [11] Sukha, S, *Advance Mathematics of Finance Honour Projects Finite-Difference Methods for Pricing the American Put Option*, 2001.
- [12] www.finance.yahoo.com/q/op?s=JPM&m=2008-02U, diakses pada tanggal 04 Februari 2008 pukul 09.45

PENGLASTERAN PASIEN KANKER LEUKEMIA BERDASARKAN DATA EKSPRESI GEN DENGAN MENGGUNAKAN DEKOMPOSISI NILAI SINGULAR

Evi Noviani¹, Yoga Satria Putra²

¹Jurusan Matematika, ²Jurusan Fisika

FMIPA Universitas Tanjungpura Kalimantan

¹evi_noviani@mipa.untan.ac.id

Abstrak

Informasi yang terkandung di dalam rangkaian molekular Deoxyribonucleic acid (DNA) makhluk hidup dapat diketahui melalui teknologi microarray. Data microarray menyajikan data tingkat ekspresi gen yang umumnya berukuran besar. Satu sampel pada data microarray bisa memiliki ribuan atau puluhan ribu gen. Pada penelitian ini diolah data pasien kanker darah (leukemia) yang berukuran 500032 dengan entri tak negatif. Data microarray pasien leukemia dapat diolah dengan menggunakan Dekomposisi Nilai Singular sedemikian sehingga sampel yang memiliki sifat yang sama dikelompokkan dalam satu kelompok. Dekomposisi Nilai Singular digunakan untuk mengelompokkan data dengan dua macam keragaman (bi-clustering), yaitu menggunakan nilai vektor singular kedua dan ketiga. Dari implementasi pada data, pasien kanker dapat dikelompokkan menjadi penyakit AML, dan ALL beserta sub tipe penyakit ALL, yakni ALL-T dan ALL-B.

Katakunci: *Data microarray, leukemia, bi-clustering, dekomposisi nilai singular.*

1. Pendahuluan

Setiap makhluk hidup, termasuk manusia, tersusun atas unit kerja terkecil yang dinamakan sel. Setiap sel yang menyusun tubuh makhluk hidup melakukan berbagai macam pekerjaan seperti metabolisme, perbanyakan diri dan lain-lain, sesuai fungsi selnya secara teratur, sehingga makhluk hidup tersebut dapat melakukan berbagai macam aktifitas. Berbagai aktifitas yang dilakukan sel telah terkodekan dalam rangkaian kode informasi genetik yang terdapat dalam rantai panjang *Deoxyribonucleic acid* (DNA). Kode informasi genetik ini mengandung bagaimana ciri-ciri, sifat, dan semua informasi lengkap tentang makhluk hidup tersebut. Pada organisme eukariotik, organisme yang memiliki membran inti sel, rangkaian DNA ini terdapat dalam inti sel.

Rangkaian informasi genetik dalam DNA pada akhirnya merupakan instruksi untuk membentuk protein tertentu. Dari perkembangan ilmu biologi dalam bidang molekuler yang semakin pesat, terdapat teknologi *microarray* yang dapat menyajikan data tingkat ekspresi gen yang dapat mengukur proses pembentukan protein ini. Satu sampel pada *microarray* bisa memiliki ribuan atau ratusan ribu gen. Sampel ini bisa berupa banyaknya orang atau organisme yang diteliti pada jaringan tertentu, atau bisa juga dari data seorang atau satu organisme saja yang diambil pada waktu yang berbeda. Sehingga data ekspresi gen yang didapatkan berupa matriks ukuran besar dengan entri dari matriks yang menunjukkan level ekspresi dari gen pada suatu sampel.

Dari data *microarray* ini perlu dianalisis sehingga dapat diketahui informasi yang terkandung di dalamnya. Salah satu informasi yang bisa didapatkan adalah pengelompokan gen atau sampel. Sampel yang mempunyai sifat yang hampir sama digolongkan dalam kelompok yang sama dan yang memiliki sifat berbeda dalam kelas yang berbeda. Proses pengelompokan ini dikenal dengan nama pengklasteran. Pengelompokan tipe-tipe kanker pada pasien sangat diperlukan sebagai pendukung diagnosa sehingga penanganannya menjadi tepat. Terdapat jenis kanker dengan ciri fisik yang sama tetapi jenisnya berbeda. Jika terdapat kekeliruan dalam diagnosa maka akan memperparah kondisi pasien. Salah satu metoda yang dapat dipakai untuk mengklasterkan data adalah menggunakan dekomposisi nilai singular yang merupakan metoda yang dipakai pada penelitian ini.

2. Landasan Teori

Deoxyribonucleic acid (DNA) merupakan pembawa informasi utama pada sebuah sel [1]. Rangkaian instruksi yang tersimpan dalam DNA mengatur aktifitas sel seperti metabolisme dan pembelahan sel. DNA ini tersusun dari molekul kecil bernama nukleotida. Terdapat empat nukleotida berbeda yang digolongkan ke dalam dua grup, yakni *purin*, terdiri dari *adenosine* dan *guanine*, serta *pyrimidin* terdiri dari *cytosine* dan *thymine*. Keempat nukleotida ini dinotasikan secara berurutan dengan *A, G, C* dan *T*. Kemudian nukleotida ini membentuk rangkaian DNA yang panjang dan membentuk struktur *double helix*.

Rangkaian DNA ini memiliki kode-kode yang merepresentasikan ciri fisik makhluk hidup, cara bekerja sel dan lain-lain. Kerja sel dapat dilihat dari protein yang dihasilkan dari kombinasi DNA ini. DNA akan mengalami transkripsi menjadi RNA (*Ribo-nucleic acid*), yaitu rangkaian nukleotida seperti halnya DNA, namun nukleotida *Thymine* diganti dengan *Uracil* (U). Pada proses transkripsi, molekul DNA pada salah satu benang disalin menjadi pre mRNA. Proses selanjutnya adalah *splicing*, yaitu membuang potongan rangkaian DNA *introns* dan menyambung rangkaian yang akan dikodekan yakni *exons* menjadi satu rangkaian mRNA. Selanjutnya dilakukan proses translasi, mRNA menuju ke ribosom yang merupakan tempat translasi. mRNA ini terbaca oleh ribosom sebagai triplet codons. Rangkaian tiga nukleotida pada DNA mengkodekan asam amino. Pada proses translasi dibentuk protein dengan menyatukan asam amino yang terkodekan pada mRNA. Protein yang dihasilkan inilah yang dapat melaksanakan fungsi yang dibutuhkan oleh sel tubuh.

Central dogma of molecular biology menyebutkan bahwa informasi tidak dapat ditransfer kembali dari protein kepada protein yang lain atau kepada asam amino. Hal ini berarti satu rangkaian DNA akan mengkodekan ekspresi gen yang menghasilkan suatu jenis protein yang unik.

Seiring perkembangan teknologi, telah ditemukan metode untuk menemukan urutan rangkaian DNA secara lengkap pada suatu makhluk hidup. Selanjutnya, perkembangan teknologi juga berhasil menemukan teknologi cDNA dan *microarray* yang dapat mengukur tingkat ekspresi gen dari puluhan ribu gen secara parallel dan menghasilkan data yang sangat besar dan berharga [2]. Ekspresi gen muncul ketika informasi molekuler tertentu

yang mengandung DNA direkam untuk menampilkan RNA-nya. Molekul inilah yang membuat protein untuk melakukan berbagai fungsi sel. Menurut *Central dogma of molecular biology*, setiap gen akan menghasilkan satu mRNA yang kemudian menghasilkan satu protein, maka level ekspresi gen satu gen dengan gen yang lain akan berbeda-beda. Teknologi *microarray* dapat mengukur ekspresi gen dalam jumlah besar secara simultan. Fragmen DNA ditempatkan dalam slide, lalu RNA yang telah diekstrak dari jaringan dan diberi label *fluorescent*, akan menyebar dalam *microarray*. Selanjutnya mengubah scan data image menjadi bentuk ekspresi gen. Seperti yang diilustrasikan [3], contoh membuat data *microarray* dari organisme ragi (*yeast*) dan [6] yang menjelaskan prosedur pembuatan data *microarray*.

Data mentah pada eksperimen *microarray* berbentuk gambar. Untuk mendapatkan data mengenai tingkat ekspresi gen, gambar tersebut harus dianalisa. Masing-masing titik diidentifikasi, diukur intensitasnya dan dibandingkan dengan latar belakangnya. Data tersebut kemudian dimasukkan kedalam sebuah hasil akhir berupa matriks ekspresi gen. Matriks inilah yang nantinya dianalisa untuk proses lebih lanjut. Pada matriks ini, baris merepresentasikan gen, dan kolom merepresentasikan berbagai macam sampel. Angka pada baris dan kolom yang bersesuaian merepresentasikan tingkat ekspresi gen tertentu pada sampel tertentu.

Data ekspresi gen disajikan dalam matriks $W_{M \times N}$ dengan N merupakan banyaknya sampel, sedangkan M adalah banyaknya gen yang diteliti. Entri dari matriks W yaitu W_{ij} merupakan level ekspresi dari gen i pada sampel j . Nilai W_{ij} yang lebih besar diantara yang lain mengandung arti level aktifitasnya lebih besar dari yang lain.

Pada penelitian ini digunakan metode Dekomposisi Nilai Singular.

Teorema 2.1 (*Dekomposisi nilai singular*) Misalkan $A \in \mathbb{R}^{(m \times n)}$ maka terdapat matriks orthogonal $U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m}$, $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$ dan matriks diagonal $\Sigma_1 \in \mathbb{R}^{r \times r}$ dengan $r \leq \min\{m, n\}$, sedemikian sehingga

$$A = U\Sigma V^T$$

Dengan Σ berukuran $m \times n$ yang mempunyai bentuk

$$\begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix}$$

Vektor-vektor kolom pada U disebut vektor singular kiri dan vektor-vektor kolom pada V disebut vektor singular kanan.

Pada penelitian ini dikelompokkan data pasien kanker darah (leukemia) [4]. Leukemia adalah penyakit yang ditandai oleh adanya akumulasi leukosit ganas dalam sumsum tulang dan darah. Sel-sel abnormal ini menyebabkan timbulnya gejala karena :

Kegagalan sumsum tulang (anemia, netropenia, trombositopenia) dan Infiltrasi organ (hati, limpa, kelenjar getah bening).

Secara tradisional Leukemia diklasifikasikan berdasarkan tipe sel yang terlihat dan bentuk maturasi sel Leukemia, yaitu:

a. Leukemia akut [8]

Leukemia akut merupakan suatu penyakit yang serius, berkembang dengan cepat, dan apabila tidak diterapi dapat menyebabkan kematian dalam beberapa minggu atau bulan. Leukemia akut dapat mempengaruhi jalan perkembangan sel limfoid akut atau jalur perkembangan sel mieolid akut [7]. Leukemia akut dapat dibagi menjadi:

1. Leukemia Limfositik Akut (LLA)

LLA adalah keganasan klonal dari sel-sel precursor limfoid. Lebih dari 80% kasus, sel-sel ganas berasal dari limfosit B dan sisanya merupakan Leukemia sel T . LLA terjadi pada 80% kasus Leukemia akut anak-anak. Insidensi puncak LLA adalah pada umur 3-7 tahun. LLA juga dapat tampak pada orang dewasa, menyebabkan 20% Leukemia akut dewasa. Tanpa pengobatan rata-rata hidup penderita LLA 3-6 bulan. Dengan pengobatan, 50% penderita LLA yang berumur antara 2-10 tahun rata-rata hidup 2-10 tahun.

2. Leukemia Granulostik/Mielositik Akut (LGA/LMA)

LGA/LMA adalah suatu penyakit yang ditandai dengan transformasi neoplastik dan gangguan diferensiasi sel-sel progenitor dari sel myeloid. Pada LGA/LMA terjadi proliferasi dari salah satu unsur sel yang memproduksi sel darah yang ganas. Sel yang ganas tersebut menginfiltrasi sumsum tulang dengan menyebabkan kegagalan fungsi tulang normal dalam proses hematopoetik normal. Dengan pengobatan angka remisi (waktu berkurangnya gejala penyakit) penderita LGA/LMA mencapai 50- 75%, tetapi angka rata-rata hidup masih 2 tahun dan yang dapat hidup lebih dari 5 tahun hanya 10%. Pada saat ini 50% anak-anak dan kira-kira 35% orang dewasa muda disembuhkan dengan kemoterapi intensif. Jika tidak ada pengobatan, penderita LGA/LMA meninggal kira-kira 3-6 bulan.

b. Leukimia kronik [5]

Leukemia kronik ditandai dengan keberadaan jumlah leukosit darah tepi yang sangat tinggi. Sel-sel ini adalah sel matur. Leukemia kronik biasanya memiliki awitan samar dan perkembangannya yang lambat. Sebagian pasien mengalami perkembangan yang lambat dan pembesaran organ yang infiltrasi oleh sel-sel Leukemia.

3. Penjelasan metodologi riset

Data ekspresi gen dari data kanker leukemia dinyatakan dalam matriks W_{MN} dengan entri $w_{ij} \geq 0$ yang merepresentasikan aktifitas gen ke- i pada sampel ke- j . Data ekspresi gen ini dapat dinyatakan menjadi graf bipartite terboboti tak negatif [6]. Titik dapat dikelompokkan ke dalam dua grup, yakni gen dan sampel, dan sisi terboboti antara keduanya ada jika pasangan titik terdapat pada grup yang berbeda (grup gen dan grup sampel). Grup gen dibagi menjadi beberapa kelompok, misalkan A, B, C, \dots . Grup sampel juga dibagi menjadi beberapa kelompok, misalkan $\hat{A}, \hat{B}, \hat{C}, \dots$. Gen di grup A cenderung aktif di sampel \hat{A} dan tidak aktif di sampel lain. Demikian juga gen di grup B aktif di grup \hat{B} dan tidak aktif di grup lain. Secara biologi, gen yang bersama dilibatkan pada fungsi tertentu akan aktif pada himpunan sampel tertentu yang memiliki kesamaan, sebagai contoh memiliki kesamaan penyakit tertentu.

Misalkan $p \in \mathbb{R}^M$ dan $q \in \mathbb{R}^N$ secara berturut-turut merupakan vektor indikator apakah suatu gen dan suatu sampel dikelompokkan pada grup

tertentu. Sampel/gen akan digolongkan menjadi kelompok-kelompok sedemikian sehingga jumlah w_{ij} terboboti yang tidak matching antara gen dan sampelnya menjadi minimum. Sehingga objektifnya adalah akan dicari p_i dan q_j yang meminimumkan resiko ketidaktepatan menempatkan gen i dan sampel j pada grup tertentu. Sehingga masalah pengklasteran dapat didekati dengan masalah minimisasi sebagai berikut:

$$\min \sum_{i=1}^M \sum_{j=1}^N (p_i - q_j)^2 w_{ij} \quad (1)$$

Kemudian dihitung jumlah total ekspresi bobot untuk setiap sampel dari gen tertentu, sebagai contoh gen i , dinotasikan $d_{gen(i)} = \sum_{k=1}^N w_{ik}$, dan jumlah total ekspresi bobot untuk setiap gen dari sampel tertentu, sebagai contoh sampel j , dinotasikan $d_{sampel(j)} = \sum_{k=1}^M w_{kj}$. Dalam bentuk matriks $W_{M \times N}$, $d_{gen(i)}$ adalah jumlah elemen dari baris i dan $d_{sampel(j)}$ adalah jumlah elemen dari kolom ke- j . Dari masing-masing jumlah total elemen baris, dibentuk vektor d_{gen} berukuran M dan vektor d_{sampel} berukuran N yang dibentuk dari masing-masing jumlah total elemen kolom. Dari vektor ini dibentuk matriks diagonal dengan elemen diagonal adalah elemen dari d_{gen} dan d_{sampel} .

$$D_{gen} = \text{diag}(d_{gen}) \in \mathbb{R}^{M \times M} \text{ dan } D_{sampel} = \text{diag}(d_{sampel}) \in \mathbb{R}^{N \times N}$$

dengan asumsi D_{gen} dan D_{sampel} matriks nonsingular.

Dengan notasi tersebut, masalah minimisasi (1) dapat ditulis

$$\min p^T D_{gen} p + q^T D_{sampel} q - 2p^T W q$$

Sehingga dapat dinyatakan dalam masalah maksimisasi:

$$\max p^T W q$$

Supaya gen dan sampel tidak hanya digolongkan ke dalam satu kelompok saja dan level ekspresi terbagi seimbang, maka ditambahkan kendala sebagai berikut:

$$\sum_{i=1}^M p_i d_{gen(i)} \approx 0 \text{ dan } \sum_{j=1}^N q_j d_{sampel(j)} \approx 0$$

Pada penelitian ini ditulis dalam bentuk perkalian vektor:

$$p^T d_{gen} = q^T d_{sampel} = 0$$

Untuk menghindari solusi trivial $p = 0$ dan $q = 0$, ditambahkan kendala normalisasi:

$$\sum_{i=1}^M p_i^2 d_{gen(i)} = 1 \text{ dan } \sum_{i=1}^N q_i^2 d_{sampel(j)} = 1 \quad (2)$$

Persamaan (2) merupakan kendala dimana jika terdapat $d_{gen(i)}$ dan $d_{sampel(j)}$ yang relatif besar (over expressed) maka akan menghasilkan p_i dan q_j yang relatif kecil, mendekati nol. Persamaan (2) dapat ditulis:

$$p^T D_{gen} p = 1 \quad (3)$$

$$q^T D_{sampel} q = 1 \quad (4)$$

Misalkan dinotasikan $D_{gen}^{\frac{1}{2}}$ adalah matriks diagonal dengan elemen diagonal $\sqrt{d_{gen(i)}}$ dan $D_{sampel}^{\frac{1}{2}}$ adalah matriks diagonal dengan elemen diagonal $\sqrt{d_{sampel(j)}}$. Maka persamaan (3) dan (4) dapat ditulis

$$\|D_{gen}^{\frac{1}{2}} p\|_2 = 1$$

$$\|D_{sampel}^{\frac{1}{2}} q\|_2 = 1$$

Sehingga masalah yang akan diselesaikan adalah berupa maksimisasi sebagai berikut:

$$\max\{p^T W q : p \in \mathbb{R}^M, q \in \mathbb{R}^N, p^T d_{gen} = q^T d_{sampel} = 0,$$

$$\|D_{gen}^{\frac{1}{2}} p\|_2 = \|D_{sampel}^{\frac{1}{2}} q\|_2 = 1\}$$

Untuk implementasi pada data, akan dilihat bagaimana SVD dapat mengelompokkan data *microarray*. Data *microarray* yang digunakan disini adalah data pasien Leukemia yang digunakan oleh [4] dimana akan dikelompokkan pasien yang termasuk acute lymphoblastic leukemia/leukemia limfositik akut (ALL) dan acute myeloid leukemia/leukemia mielositik akut (AML). Untuk kepentingan implementasi program, pada penelitian ini digunakan program matlab sehingga data dapat dibaca dengan lebih mudah melalui visualisasi.

4. Hasil diskusi

Untuk aplikasi algoritma, digunakan data *microarray* yang dipublikasikan pada [4]. Data terdiri dari 38 data gen pada jaringan tulang belakang pasien dengan ekspresi gen sebanyak 5000 gen. Dua puluh tujuh didiagnosa menderita acute lymphoblastic leukemia (ALL) dan 11 lainnya didiagnosa menderita acute myeloid leukemia (AML). Berikut daftar pasien berdasarkan data pada [4]:

Tabel 1: Diagnosa penyakit pada 38 pasien

NS	Diagnosa	NS	Diagnosa	NS	Diagnosa	NS	Diagnosa
1	<i>ALL_B</i>	11	<i>ALL_B</i>	21	<i>ALL_T</i>	31	<i>AML</i>
2	<i>ALL_B</i>	12	<i>ALL_B</i>	22	<i>ALL_T</i>	32	<i>AML</i>
3	<i>ALL_B</i>	13	<i>ALL_B</i>	23	<i>ALL_T</i>	33	<i>AML</i>
4	<i>ALL_B</i>	14	<i>ALL_B</i>	24	<i>ALL_T</i>	34	<i>AML</i>
5	<i>ALL_B</i>	15	<i>ALL_B</i>	25	<i>ALL_T</i>	35	<i>AML</i>
6	<i>ALL_B</i>	16	<i>ALL_B</i>	26	<i>ALL_T</i>	36	<i>AML</i>
7	<i>ALL_B</i>	17	<i>ALL_B</i>	27	<i>ALL_T</i>	37	<i>AML</i>
8	<i>ALL_B</i>	18	<i>ALL_B</i>	28	<i>AML</i>	38	<i>AML</i>
9	<i>ALL_B</i>	19	<i>ALL_B</i>	29	<i>AML</i>		
10	<i>ALL_B</i>	20	<i>ALL_T</i>	30	<i>AML</i>		

NS: Nomor sampel.

Dari data awal, pertama kolom data diacak dengan tetap mengingat nomor sampel awal. Berikut penomoran baru yang digunakan:

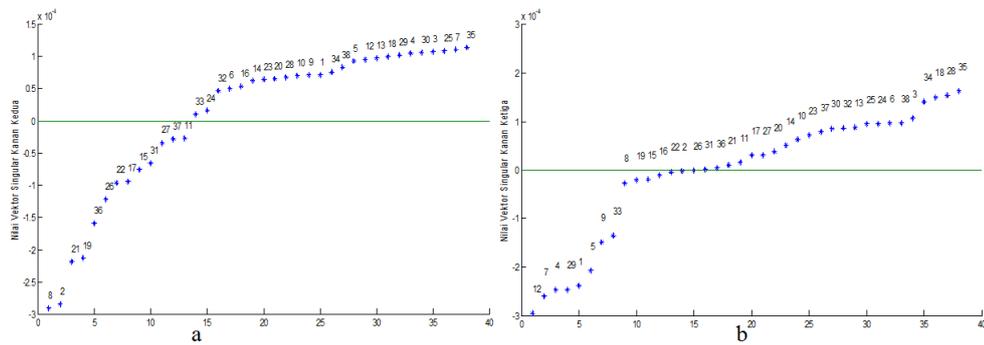
Tabel 2: Diagnosa penyakit pada 38 pasien

NS	1	2	3	4	5	6	7	8	9	10	11	12	13
NB	38	10	23	20	32	27	18	3	30	16	24	28	35
NS	14	15	16	17	18	19	20	21	22	23	24	25	26
NB	25	6	13	37	34	14	33	4	12	29	1	5	9
NS	27	28	29	30	31	32	33	34	35	36	37	38	
NB	7	22	11	8	2	26	17	36	19	15	31	21	

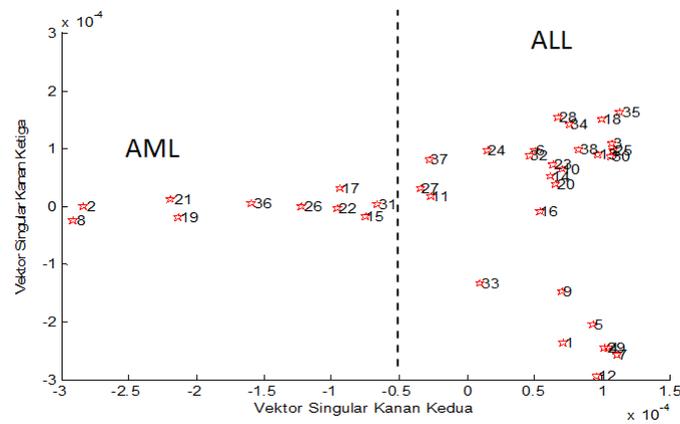
NS: Nomor sampel, NB: Nomor baru.

Berikut pada Gambar 1(a) diplot nilai $D_{sampel}^{-\frac{1}{2}}v^{(2)}$ dengan $v^{(2)}$ merupakan vektor singular kanan kedua dari matriks $D_{gen}^{-\frac{1}{2}}WD_{sampel}^{-\frac{1}{2}}$ setelah diurutkan dari yang terkecil ke yang terbesar. Angka yang terletak di atas titik (*) menunjukkan nomor sampel baru hasil pengacakan. Dari gambar tersebut dapat terlihat adanya dua pengelompokan sampel, yakni yang memiliki nilai vektor singular kanan yang positif dan negatif. Nomor sampel {8, 2, 21, 19, 36, 26, 22, 17, 15, 31, 27, 37, 11} memiliki nilai singular vektor kanan yang kedua yang negatif. Hal ini bersesuaian dengan nomor sampel yang didiagnosa memiliki penyakit AML (lihat tabel 1 dan 2), kecuali nomor sampel 27 dan 37, yang bersesuaian dengan nomor sampel awal 6 dan 17 sebelumnya didiagnosa memiliki penyakit ALL-B. Ketidaksesuaian ini bisa disebabkan karena ketidakkonsistenan data, seperti pembulatan atau pada saat pemilihan gen pada data yang tidak melibatkan semua gen (dalam contoh ini hanya 5000 gen). Ketidaksesuaian ini juga bisa disebabkan kesalahan pada diagnosa sehingga diperlukan metode lebih lanjut untuk menelitinya. Sedangkan sisanya memiliki nilai vektor singular kanan kedua yang positif. Hal ini bersesuaian dengan nomor sampel yang didiagnosa memiliki penyakit ALL. Dari Gambar 1(a) ini terlihat bahwa vektor singular kanan kedua dapat mengelompokkan pasien dengan diagnosa penyakit AML dan ALL. Berikut pada Gambar 1(b) diplot nilai $D_{sampel}^{-\frac{1}{2}}v^{(3)}$ dengan $v^{(3)}$ merupakan vektor singular kanan ketiga dari matriks $D_{gen}^{-\frac{1}{2}}WD_{sampel}^{-\frac{1}{2}}$ setelah diurutkan dari yang terkecil ke yang terbesar.

Dari Gambar 1(b) dapat terlihat adanya pengelompokan sampel. Sampel dengan nomor 12, 7, 4, 29, 1, 5, 9, dan 33 tergolong kedalam satu kelompok yang sama, sebut sebagai kelompok A dan nomor sampel yang lain (8, 19, 15, 16, 22, 2, 26, 31, 36, 21, 11, 17, 27, 20, 14, 10, 23, 37, 30, 32, 13, 25, 24, 6, 38, 3, 34, 18, 28 dan 36) ke dalam kelompok lain, sebut kelompok B. Sampel yang termasuk kelompok A bersesuaian dengan sampel yang didiagnosa memiliki penyakit ALL-T, sedangkan sampel yang termasuk kelompok B didiagnosa memiliki penyakit ALL-B atau memiliki penyakit AML (lihat Tabel 1 dan 2). Dari gambar ini dapat terlihat bahwa nilai vektor singular kanan ketiga dapat mengelompokkan secara nyata sub penyakit ALL, yakni ALL-T dari ALL-B dan AML. Jika diplot antara nilai



Gambar 1: Nilai dari vektor singular kanan kedua (a) dan ketiga (b) setelah pengurutan



Gambar 2: Pengelompokan sampel berdasarkan vektor singular kanan kedua dan ketiga

vektor singular kanan kedua dan ketiga maka akan menghasilkan pengelompokan seperti pada Gambar 2. Pada gambar tersebut dapat terlihat adanya pengelompokan sampel pasien menjadi tiga kelompok. Pada perhitungan ini, nilai singular kesatu sampai keempat yang didapatkan secara berturut-turut adalah 1, 0.38, 0.296 dan 0.234. Adapun persentase kontribusinya jika digunakan singular kedua dan ketiga adalah 13.7. Dengan nilai kontribusi tersebut dapat dikelompokkan data sampel dengan penyakit AML, dan ALL beserta sub tipe penyakit ALL, yakni ALL-T dan ALL-B meski

pengelompokannya tidak secara nyata, yakni jarak antar kelompok tidak terpisah jauh.

5. Penutup

Berdasarkan hasil dan pembahasan yang telah dijelaskan sebelumnya, maka dapat diambil beberapa kesimpulan bahwa dekomposisi nilai singular dapat digunakan untuk mengelompokkan data dengan ukuran relatif besar dalam hal ini data pasien kanker leukemia. Data *microarray* dengan entri seluruhnya tak negatif dapat dikelompokkan dengan menggunakan nilai singular kedua sebagai hasil dekomposisi nilai singular yang diterapkan pada data dengan menggunakan normalisasi terlebih dahulu. Dan untuk kepentingan visualisasi dapat disertakan juga nilai vektor singular berikutnya (kedua atau ketiga). Dari implementasi pada data, pasien kanker dapat dikelompokkan menjadi penyakit AML, dan ALL beserta sub tipe penyakit ALL, yakni ALL-T dan ALL-B.

Pustaka

- [1] BRAZMA, ALVIS, HELEN P., THOMAS S., MOHAMMADREZA S., *A quick introduction to elements of biology - cells, molecules, genes, functional genomics, microarrays*, European Bioinformatics Institute, on-line tutorial di www.ebi.ac.uk/microarray/biology_intro.html, 2001.
- [2] BRAZMA, ALVIS. AND JAAK VILO, *Gene Expression Data Analysis*, Federation of European Biochemical Societies:480. p. 17-24, 2000.
- [3] BROWN, PATRICK O. AND DAVID BOTSTEINS, . *Exploring the new world of the genome with DNA microarrays*, Nature genetics supplement: 21, p. 33-37, 1999.
- [4] BRUNET, JP., PABLO TAMAYO, T.R. GOLUB, AND J.P. MESIROV, *Metagenes and Molecular Pattern Discovery Using Matrix Factorization.*, Procidings of The National Academy of Sciences: 101. p. 4164-4169, 2004

-
- [5] CHANDRASOMA, P, TAYLOR, CLIVE R., *Ringkasan Patologi Anatomi*, Edisi ke-2. Jakarta:EGC, 2005.
- [6] CHEUNG, VIVIAN G , MICHAEL MORLEY, FRANCISCO AGUILAR, ALDO MASSIMI,RAJU KUCHERLAPATI AND GEOFFREY CHILDS, *Making and reading microarrays*, Nature genetics supplement: 21, p. 15-19, 1999
- [7] RAHMALIA, A, NOVIANTI R., *At a Glance Medicine*, Jakarta: Erlangga., 2006.
- [8] SETIAWAN, L,*Kapita Selekt Hematologi*, Edisi Keempat. Jakarta: EGC, 2005.

ESTIMASI MODEL EKSPONENSIAL *LIFETIME* DENGAN *DOUBLE CENSORING*

Farida Agustini W.¹, Thathit Purwaningtyas²

Jurusan Matematika, FMIPA ITS Surabaya

¹farida_sahlan@yahoo.com, ²thathit@sss_sub.sig.co.id

Abstrak

Analisis *lifetime* merupakan analisis statistik yang memodelkan lama waktu sampai terjadinya suatu kejadian. Penelitian ini menggunakan Distribusi Eksponensial untuk mengestimasi data *lifetime* dengan *double censoring* dan menggunakan metode Bayesian yang berdasarkan distribusi prior gamma. Informasi sampel (distribusi prior) tersebut dikombinasikan dengan fungsi *likelihood* untuk mendapatkan distribusi posterior, sehingga didapatkan estimator dari μ yaitu $\bar{\mu}$ dan interval kepercayaan $(1 - \alpha) 100\%$ untuk μ , yaitu: $P(c_1 \leq \bar{\mu} \leq c_2 | x) = 1 - \alpha$ dimana penentuan nilai (c_1, c_2) didapatkan dengan pendekatan numerik dengan bantuan Matlab 7.0 pada penerapan kasus.

Katakunci: *Distribusi eksponensial, double censoring, metode bayesian, distribusi posterior*

1. Pendahuluan

Pengamatan data *lifetime* berguna dalam melakukan pengujian daya tahan dan keandalan suatu hasil teknologi dan industri. Selain sangat diperlukan dalam peningkatan kualitas suatu produk, analisis terhadap data *lifetime*

pada kondisi perlakuan atau operasional tertentu banyak dilakukan oleh para ilmuwan di berbagai bidang ilmu.

Beberapa model distribusi yang sering digunakan dalam pengamatan data *lifetime* ini yaitu distribusi Weibull, distribusi Gamma, distribusi Eksponensial dan distribusi Lognormal. Pada penelitian ini distribusi yang digunakan adalah distribusi Eksponensial karena distribusi tersebut banyak digunakan dalam analisis uji *lifetime*. Pengambilan sampel dalam uji *lifetime* dilakukan secara *censoring* sehingga waktu pengujian lebih efisien. Terdapat berbagai macam tipe sampel *censoring* yang bisa digunakan diantaranya sampel *single censoring* dan *double censoring*. Sampel *double censoring* adalah sampel data yang disensor kedua-duanya yaitu *right censoring* dan *left censoring*, sedangkan sampel *single censoring* merupakan tipe *censoring* yang hanya dilakukan satu kali, yaitu *right censoring* saja atau *left censoring*. Penggunaan sampel *double censoring* lebih menghemat waktu dan tenaga jika dibandingkan dengan sampel *single censoring* karena pengamatan bisa dilakukan dengan menggunakan data yang tersedia. Penelitian ini meliputi estimasi parameter yang menggunakan Metode *Bayes*, yang merupakan metode estimasi berdasarkan penggabungan informasi data sampel dengan pengetahuan subyektif mengenai distribusi peluang parameter yang tidak diketahui yang disebut distribusi prior dan prior yang digunakan adalah prior *inverse* Gamma.

Tujuannya dari penelitian ini adalah untuk mendapatkan estimator parameter distribusi Eksponensial pada data *double censoring* berdasarkan distribusi prior *inverse* Gamma serta interval kepercayaannya.

2. Dasar teori

Misalkan $G(x)$ adalah fungsi dari variabel acak x , nilai harapan untuk $G(x)$ adalah:

$$E[G(x)] = \sum_{x \in \mathbb{R}} G(x)f(x) \text{ jika } x \text{ diskrit}$$

$$E[G(x)] = \int_{-\infty}^{\infty} G(x)f(x) dx \text{ jika } x \text{ kontinu}$$

Misalkan X dan Y adalah distribusi bersama dari variabel acak maka nilai harapan bersyarat dari Y dimana $X = x$ adalah :

$$E[Y|(x)] = \sum_y yf(y|x) \text{ jika } X \text{ dan } Y \text{ diskrit}$$

$$E[Y|(x)] = \int_{-\infty}^{\infty} yf(y|x) dx \text{ jika } X \text{ dan } Y \text{ kontinu}$$

Definisi 2.1 [3] Sebuah variabel acak kontinu X dikatakan memiliki **distribusi Inverse Gamma** dengan parameter bentuk α dan parameter skala β , jika fungsi kepadatan probabilitas (pdf) dari X berbentuk

$$f(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha-1)} e^{-\frac{\beta}{x}} & ; \text{ untuk } x > 0 \\ 0 & ; \text{ untuk } x \text{ lainnya} \end{cases} \text{ dan } \alpha > 0 ; \beta > 0$$

Dalam analisis Bayesian biasanya keputusan penentuan distribusi prior memegang peranan penting. Dengan menggunakan teorema Bayes, informasi awal (prior) ini digunakan untuk mendapatkan nilai dari estimasi μ , yang dinyatakan dengan fungsi *likelihood* yang dikombinasikan untuk membentuk distribusi posterior. Distribusi Prior dari parameter μ dilambangkan dengan $p(\mu)$ adalah fungsi probabilitas yang menyatakan tingkat kepercayaan nilai μ , distribusi posterior $\propto \text{likelihood} \times$ distribusi prior.

Beberapa *censoring* dalam analisis uji hidup adalah [11]:

a. *Uncensored*

Dalam uji sampel lengkap ini eksperimen akan dihentikan jika semua komponen yang diuji telah mati atau gagal.

b. *Left Censoring*

Jika sebuah titik point berada di bawah sebuah nilai tertentu tetapi tidak diketahui berapa banyaknya.

c. *Right Censoring*

Jika sebuah titik point berada diatas sebuah nilai tertentu tetapi tidak diketahui berapa banyaknya [2].

Right Censoring terdiri dari:

- i *Censoring* Tipe I: Dalam censoring tipe I, eksperimen akan dihentikan jika telah dicapai waktu tertentu (waktu censoring).
- ii *Censoring* Tipe II: Suatu sampel dikatakan censoring tipe II apabila eksperimen dihentikan setelah kegagalan ke- r telah diperoleh..

d. *Double Censoring*[5]

Dalam Double Censoring, pengamatan terhadap waktu hidup n sampel acak dengan dua censoring yaitu *right censoring* dan *left censoring*, sedangkan untuk yang *right censoring* digunakan *right censoring* tipe II. *Double censoring* dari n sample acak adalah pengamatan terhadap waktu hidup sampel ke X_r, X_{r+1}, \dots, X_s , yang terurut dengan $(r-1)$ sampel terkecil serta $(n-s)$ sampel terbesar disensor.

$$\overbrace{X_1, X_2, \dots, X_{r-1}}^{\text{disensor}}, \underbrace{X_r, X_{r+1}, \dots, X_s}_{\text{diamati}}, \overbrace{X_{s+1}, X_{s+2}, \dots, X_n}^{\text{disensor}}$$

Salah satu metode yang digunakan untuk mendapatkan estimator dari parameter adalah Metode *Bayesian Estimation*. Parameter μ dipandang sebagai variabel acak dalam ruang parameter μ dan mempunyai distribusi probabilitas $p(\mu)$ yang merupakan tingkat kepercayaan awal tentang parameter μ sebelum pengamatan dilakukan, disebut distribusi prior μ . Teorema umum dari Bayes [4]

$$p(\mu|y) = \frac{p(y|\mu)p(\mu)}{p(y)}$$

dimana $p(\mu|y)$ distribusi posterior μ . Distribusi bersama $p(y, \mu)$ dan distribusi marginal $p(y)$ pada umumnya tidak diketahui, biasanya hanya distribusi prior dan fungsi *likelihood*nya yang dinyatakan. Rumus Bayes dapat juga ditulis [4]

$$p(\mu|y) \propto l(\mu, y) p(\mu)$$

Distribusi posterior \propto *likelihood* \times distribusi prior

Interval (c_1, c_2) dari $100(1 - \alpha)\%$ disebut Interval Bayesian untuk μ jika $c_1 < c_2$ adalah dua konstanta, maka [7]:

$$1 - \alpha \leq P(C|x) = \begin{cases} \int_C \pi(\theta|x)d\theta, & \text{kontinu;} \\ \sum_{\theta \in C} \pi(\theta|x), & \text{diskrit.} \end{cases}$$

dengan $\pi(\theta|x)$ adalah distribusi posterior.

3. Analisis dan pembahasan

3.1. Estimasi parameter distribusi eksponensial

Distribusi Eksponensial secara luas digunakan sebagai suatu model *lifetime* yang didasarkan penelitian yang menyertakan ketahanan (*survival*). Jika x *lifetime* berdistribusi Eksponensial dengan parameter μ maka pdf $f(x)$ adalah [6]

$$f(x) = \frac{1}{\mu} \exp\left(-\frac{x}{\mu}\right) \quad x > 0 \quad \mu > 0$$

3.2. Fungsi survival

Probabilitas suatu individu oleh distribusi Eksponensial yang akan bertahan hidup sampai waktu x disebut fungsi *survival* adalah:

$$R(x|\mu) = \Pr(X > x|\mu) \quad (1)$$

Sehingga dari Persamaan (1) didapatkan *fungsi survival* sebagai berikut:

$$R(x|\mu) = \exp\left(-\frac{x}{\mu}\right) \quad (2)$$

3.3. Fungsi Likelihood

Misalkan x variabel acak berdistribusi Eksponensial dengan fungsi probabilitas $f(x, \mu)$, dan μ adalah parameter yang tidak diketahui maka pengkonstruksian fungsi *likelihood* dapat dinyatakan dengan

$$\begin{aligned}
L(x|\mu) &= f(x_1, \mu)f(x_2, \mu)\dots f(x_n, \mu) \\
&= \prod_{i=1}^n f(x_i; \mu)
\end{aligned} \tag{3}$$

Fungsi *likelihood* pada Persamaan (2) berlaku untuk tipe *complete sampling*, karena data pada tipe ini semua *uncensored*, artinya semua unit yang diuji gagal. Sedangkan untuk tipe *censored sampling* mempunyai fungsi *likelihood* yang berbeda dengan *complete sampling*. Ketika mengerjakan analisa *maximum likelihood* pada data dengan item yang belum gagal, fungsi *likelihood* perlu diperluas untuk menghitung item-item yang belum gagal. Secara keseluruhan teknik estimasinya tidak berubah, hanya saja pada fungsi *likelihood* dikenakan fungsi *survivor*. Fungsi *likelihood* untuk model *survival* dengan keadaan data *censoring*, diberikan formulasi berikut ini [8]:

i Untuk *uncensored*: $\Pr(X = x|\mu) = f(x|\mu)$

Uji pada sampel lengkap ini eksperimen akan dihentikan jika semua komponen yang diuji telah mati atau gagal.

ii Untuk *left censoring*: $\Pr(X < x|\mu) = 1 - R(x_r|\mu)$

Jika sebuah titik point berada di bawah sebuah nilai tertentu tetapi tidak diketahui berapa banyaknya.

iii Untuk *right censoring*: $\Pr(X > x|\mu) = R(x_r|\mu)$

Jika sebuah titik point berada diatas sebuah nilai tertentu tetapi tidak diketahui berapa banyaknya.

Sample acak berukuran n dari distribusi $\text{Exp}(\mu)$, dimana $\mu \in U = (0, \infty)$ adalah parameter yang tidak diketahui, dan menurut Fernandez (2000) misalkan x_r, \dots, x_s merupakan pengamatan terurut dari $(r - 1)$ pengamatan terkecil dan $(n - s)$ pengamatan terbesar pada *double censoring* adalah sebagai berikut:

$$\overbrace{x_1, x_2, \dots, x_{r-1}}^{\text{disensor}}, \underbrace{x_r, x_{r+1}, \dots, x_s}_{\text{diamati}}, \overbrace{x_{s+1}, x_{s+2}, \dots, x_n}^{\text{disensor}}$$

selanjutnya substitusi Persamaan (2) dan (3) didapatkan fungsi *likelihood* untuk μ pada data *double censoring* adalah

$$L(\mu|x) = \frac{n!}{(r-1)!(n-s)!} \{1 - R(x_r|\mu)\}^{r-1} \{R(x_r|\mu)\}^{n-s} \prod_{i=r}^s f(x_i|\mu) \quad (4)$$

sehingga diperoleh

$$\begin{aligned} L(\mu|x) &= \frac{n!}{(r-1)!(n-s)!} \left\{ 1 - \exp\left(-\frac{x_r}{\mu}\right) \right\}^{r-1} \\ &\quad \times \left\{ \mu^{-(s-r+1)} \exp\left(-\frac{\sum_{i=r}^s x_i + (n-s)x_s}{\mu}\right) \right\} \end{aligned} \quad (5)$$

Pada Persamaan (4) dimisalkan $m = s - r + 1$ dan $\xi(x) = \sum_{i=r}^s x_i + (n-s)x_s$ sehingga

$$L(\mu|x) = \frac{n!}{(r-1)!(n-s)!} \left\{ 1 - \exp\left(-\frac{x_r}{\mu}\right) \right\}^{r-1} \left\{ \mu^{-m} \exp\left(-\frac{\xi(x)}{\mu}\right) \right\} \quad (6)$$

dengan kesebandingan didapatkan fungsi *likelihood* untuk μ adalah

$$L(\mu|x) \propto \left\{ 1 - \exp\left(-\frac{x_r}{\mu}\right) \right\}^{r-1} \left\{ \mu^{-m} \exp\left(-\frac{\xi(x)}{\mu}\right) \right\} \quad (7)$$

3.4. Distribusi prior

Menurut [5] distribusi yang tepat untuk normasi prior dari μ adalah distribusi *Inverse Gamma* (a, b) . Distribusi prior tersebut berbentuk

$$g(\mu) = \frac{a^b}{\Gamma(b)} \mu^{-(b+1)} \exp\left(-\frac{a}{\mu}\right)$$

dimana $\mu > 0$, $a > 0$, $b > 0$. dengan kesebandingan didapatkan distribusi prior sebagai berikut:

$$g(\mu) \propto \mu^{-(b+1)} \exp\left(-\frac{a}{\mu}\right) \quad \text{dimana } \mu > 0, a > 0, b > 0 \quad (8)$$

dengan a dan b diketahui.

3.5. Distribusi posterior

Metode Bayesian mempunyai keuntungan dengan mengkombinasikan pengetahuan subjektif dari distribusi prior dengan pengetahuan yang ada dalam data.

Dengan mengkombinasikan fungsi *likelihood* dan distribusi prior didapatkan distribusi posteriornya. Persamaan umum dari distribusi posterior tersebut adalah :

$$h(\mu|x) = \frac{L(\mu|x)p(\mu)}{\int L(\mu|x)p(\mu)d\mu} \quad (9)$$

dengan $L(\mu|x)$ adalah fungsi *likelihood* dan $p(\mu)$ adalah distribusi prior.

Berdasarkan Persamaan (6) dan (8) nilai dari $\int L(\mu|x)p(\mu)d\mu$ adalah:

$$\begin{aligned} & \int L(\mu|x)p(\mu)d\mu \\ &= \int_0^{\infty} \mu^{-(b+1)} \exp\left(\frac{-a}{\mu}\right) \mu^{-m} \exp\left(\frac{-\xi(x)}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu \\ &= \int_0^{\infty} \mu^{-(m+b+1)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu \quad (10) \end{aligned}$$

dengan

$$\begin{aligned} \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} &= \sum_{i=0}^{r-1} \binom{r-1}{i} \left(-\exp\left(-\frac{x_r}{\mu}\right)\right)^i (1)^{r-1-i} \\ &= \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \exp\left(-\frac{ix_r}{\mu}\right) \quad (11) \end{aligned}$$

Kemudian substitusikan Persamaan (10) ke (11), sehingga didapatkan

$$\begin{aligned}
& \int L(\mu|x) p(\mu) d\mu \\
&= \int_0^{\infty} \mu^{-(m+b+1)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \exp\left(-\frac{ix_r}{\mu}\right) d\mu \\
&= \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \int_0^{\infty} \mu^{-(m+b+1)} \exp\left(\frac{-(a+\xi(x)+ix_r)}{\mu}\right) d\mu \\
&= \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \frac{\Gamma(-b+m)}{(a+\xi(x)+ix_r)^{(b+m)}} \\
& \int_0^{\infty} \frac{(a+\xi(x)+ix_r)^{(b+m)}}{\Gamma(b+m)} \mu^{-(m+b+1)} \exp\left(\frac{-(a+\xi(x)+ix_r)}{\mu}\right) d\mu \quad (12)
\end{aligned}$$

maka $\int_0^{\infty} \frac{(a+\xi(x)+ix_r)^{(b+m)}}{\Gamma(b+m)} \mu^{-(m+b+1)} \exp\left(\frac{-(a+\xi(x)+ix_r)}{\mu}\right) d\mu = 1$

sehingga Persamaan (12) akan menjadi

$$\int L(\mu|x) p(\mu) d\mu = \Gamma(b+m) \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \left(\partial \left(1 + \frac{ix_r}{\partial}\right)\right)^{-(b+m)}$$

dengan $\partial : a + \xi(x)$, diperoleh distribusi posteriornya dinyatakan sebagai berikut:

$$h(\mu|x) = \frac{\mu^{-(b+1)} \exp\left(\frac{-a}{\mu}\right) \mu^{-m} \exp\left(\frac{-\xi(x)}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu}{\Gamma(b+m) \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^{-i} \left(a + \xi(x) \left(1 + \frac{ix_r}{a + \xi(x)}\right)\right)^{-(b+m)}} \quad (13)$$

Misalkan

$$Fr[a + \xi(x), b + m] = \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^{-i} \left(1 + \frac{ix}{a + \xi(x)}\right)^{-(b+m)}$$

dengan $a + \xi(x), m + b > 0$. Sehingga distribusi posterior pada Persamaan

(13) dapat dinyatakan sebagai berikut:

$$h(\mu|x) = \frac{\mu^{-(b+m+1)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu}{\Gamma(b+m) (a+\xi(x))^{-(b+m)} Fr[a+\xi(x), b+m]} \quad (14)$$

3.6. Estimator Bayes

Dengan menggunakan fungsi kerugian kuadratik, estimator Bayes untuk $\mu(\bar{\mu})$ adalah mean dari distribusi posterior. Estimasi Bayes dapat dinyatakan sebagai berikut:

$$\bar{\mu} = E(\mu|x) = \int_0^{\infty} \mu h(\mu|x) d\mu \quad (15)$$

dimana $h(\mu|x)$ adalah distribusi posterior sesuai Persamaan (14), sehingga Persamaan (15) menjadi

$$\begin{aligned} \bar{\mu} &= \int_0^{\infty} \mu \frac{(a+\xi(x))^{(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu}{\Gamma(b+m) \mu^{(b+m+1)} Fr[a+\xi(x), b+m]} \\ &= \frac{(a+\xi(x))^{(b+m)}}{\Gamma(b+m) Fr[a+\xi(x), b+m]} \\ &\quad \int_0^{\infty} \mu^{-(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu \quad (16) \end{aligned}$$

Persamaan $\int_0^{\infty} \mu^{-(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu$ dapat ditulis sebagai berikut:

$$\int_0^{\infty} \mu^{-(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu$$

karena

$$\int_0^{\infty} \frac{(a+\xi(x) + ix_r)^{(b+m-1)}}{\Gamma(b+m+1)} \mu^{-((b+m-1)+1)} \exp\left(\frac{-(a+\xi(x) + ix_r)}{\mu}\right) d\mu = 1$$

sehingga

$$\begin{aligned}
& \int_0^{\infty} \mu^{-(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu \\
&= \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \frac{\Gamma(b+m-1)}{(a+\xi(x)+ix_r)^{(b+m-1)}} \\
&= \Gamma(b+m-1) (\partial(x))^{-(b+m-1)} \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^i \left(1 + \frac{ix_r}{\partial}\right)^{-(b+m-1)} \quad (17)
\end{aligned}$$

dengan $\partial : a + \xi(x)$.

Misalkan

$$Fr[a + \xi(x), b + m - 1] = \sum_{i=0}^{r-1} \binom{r-1}{i} (-1)^{-i} \left(1 + \frac{ix}{a + \xi(x)}\right)^{-(b+m-1)}$$

dengan $a + \xi(x), m + b - 1 > 0$

Sehingga Persamaan (17) dapat ditulis sebagai berikut:

$$\begin{aligned}
& \int_0^{\infty} \mu^{-(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) \left(1 - \exp\left(-\frac{x_r}{\mu}\right)\right)^{r-1} d\mu \\
&= \Gamma(b+m+1) (a + \xi(x))^{-(b+m-1)} Fr[a + \xi(x), b + m - 1] \quad (18)
\end{aligned}$$

Dengan demikian estimator Bayes untuk μ pada Persamaan (16) dan substitusi dari Persamaan (18) dapat dinyatakan sebagai berikut:

$$\bar{\mu} = \frac{\Gamma(b+m-1) Fr[a + \xi(x), b + m - 1]}{\Gamma(b+m) Fr[a + \xi(x), b + m]} (a + \xi(x))$$

dengan $\Gamma(b+m) = (b+m-1) \Gamma(b+m-1)$, maka estimator Bayes untuk μ adalah sebagai berikut:

$$\bar{\mu} = \left\{ \frac{Fr[a + \xi(x), b + m - 1]}{Fr[a + \xi(x), b + m]} \right\} \left\{ \frac{(a + \xi(x))}{b + m - 1} \right\}$$

dengan $m = s - r - 1$ dan $\xi(x) = \sum_{i=r}^s x_i + (n - s) x_s$

3.7. Interval kepercayaan Bayes

Diketahui bahwa distribusi posterior dari μ adalah:

$$h(\mu|x) = \frac{(a + \xi(x))^{(b+m)} \exp\left(\frac{-(a+\xi(x))}{\mu}\right) (1 - \exp(-\frac{x}{\mu}))^{r-1}}{\Gamma(b + m)\mu^{(b+m+1)} Fr[a + \xi(x), b + m]},$$

dengan $\mu > 0$ yaitu $100(1 - \alpha)\%$ interval kepercayaan untuk μ dan $(1 - \alpha)$ disebut koefisien kepercayaan

$$1 - \alpha \leq P(C|x) = \begin{cases} \int_C \pi(\theta|x) d\theta, & \text{kontinu;} \\ \sum_{\theta \in C} \pi(\theta|x), & \text{diskrit.} \end{cases}$$

dengan $\pi(\theta|x)$ adalah distribusi posterior.

Selanjutnya akan dicari interval kepercayaan $(1 - \alpha)100\%$ untuk μ , yaitu: $P(c_1 \leq \mu \leq c_2|x) = 1 - \alpha$ dimana μ adalah parameter yang tidak diketahui dengan batas bawah memenuhi persamaan

$$S(c_1|x) = \int_0^{c_1} h(\mu|x) d\mu = \frac{\alpha}{2}$$

dan batas atas juga memenuhi persamaan

$$S(c_2|x) = \int_0^{c_2} h(\mu|x) d\mu = 1 - \frac{\alpha}{2}$$

Batas bawah (c_1) dan batas atas (c_2) ditentukan dengan menggunakan pendekatan numerik. Pendekatan integrasi numerik yang digunakan pada penelitian adalah Aturan Simpson's $\frac{1}{3}$ berdasarkan dengan bantuan matlab 7.0.

Contoh kasus Ilustrasi permasalahan *flure time* (waktu kegagalan) dengan satuan menit, untuk jenis sekatan listrik type B pada suatu percobaan dimana sekatan digunakan terus menerus untuk meningkatkan tekanan voltase.

..., 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, ...

Pada contoh ini pengamat gagal untuk mengamati dua waktu kegagalan paling kecil dan pengamatan dihentikan pada saat waktu kegagalan ke-9

dan diketahui $a = 26$ dan $b = 27$. Oleh karena itu didapatkan $n = 12, r = 3, s = 9$, dan $m = 7$. sehingga diperoleh: $\bar{\mu} = 20,47555772$ dan interval kepercayaannya

$$15.75 < \mu < 28.75$$

4. Kesimpulan dan saran

4.1. Kesimpulan

Berdasarkan hasil dan pembahasan disimpulkan sebagai berikut:

- a. Estimator parameter distribusi Eksponensial pada data *double censoring* berdasarkan prior *inverse gamma* adalah:

$$\bar{\mu} = \left\{ \frac{Fr [a + \xi(x), b + m - 1]}{Fr [a + \xi(x), b + m]} \right\} \left\{ \frac{(a + \xi(x))}{b + m - 1} \right\}$$

- b. Interval kepercayaan $(1 - \alpha)$ 100% parameter distribusi Eksponensial pada data *Double censoring* berdasarkan distribusi prior *inverse Gamma* dengan $\alpha = 0,05$, yaitu

$$P(c_1 \leq \mu \leq c_2 | x) = 1 - \alpha.$$

dengan batas = 0,025 dan batas atas = 0,975.

Nilai batas bawah (c_1) dan batas atas (c_2) didapatkan dengan pendekatan numerik Aturan Simpson's $\frac{1}{3}$ dengan bantuan Matlab 7.0. Pada penerapan contoh kasus didapatkan nilai estimasi ($\bar{\mu}$) = 20,47555772 dengan $\alpha = 0,05$ dan didapatkan interval kepercayaan bayes dengan nilai batas bawah (c_1) = 15.75 dan batas atas (c_2) = 28.75 sehingga interval kepercayaannya $15.75 < \mu < 28.75$

4.2. Saran

Saran yang dapat diberikan pada pembahasan estimasi Eksponensial *lifetime* dengan *double censoring* yang menggunakan metode Estimasi Bayesian dan berdasarkan distribusi prior *inverse gamma*, dapat dikembangkan lebih lanjut dengan pendekatan distribusi prior yang lain misalnya : *Non-Informatif Prior, Conjugate Prior, Prior Proper, Prior Improper* serta *Reference Prior*.

Pustaka

- [1] Bain, L. J. dan Max E., *Introduction to Probability and Mathematical Statistics*, 2nd edition, California: Duxbury Press, 1991.
- [2] Berger, J.O., *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. , Springer – Verlag, New York, 1998.
- [3] Bernardo, J.M, dan Smith, A.F.M., *Bayesian theory*, Chichester : John Wiley & Sons, Ltd, 2000.
- [4] Box, G., dan Tiao, G. C., *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Company, Inc., Philippines, 1973
- [5] Fernandez, J. Arturo., *Estimation and Hypothesis Testing for Eksponential Lifetime Models with Double censoring and Prior Information*, Journal of Economic and Social Research, 2000.
- [6] Harinaldi., *Prinsip-Prinsip Statistik Untuk Teknik dan Sains*, Erlangga, Jakarta, 2006.
- [7] Larson, H. J., *Introduction to Probability Theory and Statistical Inference* , John Wiley and Sons, Singapore, 1982.
- [8] Miller, R. G., *Survival Analysis* , Wiley classics library, Stanford University, Canada, 1998.
- [9] Sarwoko, *Statistik Inferensi untuk Ekonomi dan Bisnis*, Andi, Yogyakarta, 2007.
- [10] Soehardjo, *Analisis Numerik*, Dosen Matematika ITS dan Institut Adhi Tama Surabaya, Surabaya, 1986.
- [11] Wikipedia, the free encyclopedia., *Types Of Censoring*, <http://www.wikipedia.com>., Diakses pada tanggal 23 Juni 2008.

CRITICAL SET OF EDGE MAGIC TOTAL LABELING OF EXPANDING CYCLE GRAPH*

Chairul Imron¹ and Suhud Wahyudi²

Mathematics Department, FMIPA ITS Surabaya

¹imron-its@matematika.its.ac.id, ²suhud@matematika.its.ac.id

Abstract

This paper discusses about graph labeling. We will find edge magic total labeling of cycle and expanding cycle graph. In the final, we investigate the critical set of edge magic total labeling on cycle and expanding cycle graph.

Keyword: *Cycle graph, edge-magic total labeling, critical set*

1. Introduction

Let $G = (V, E)$ be a finite simple and undirected graphs. That is to say, they contain neither loops nor multiple edges. The graph have vertex-set $V(G)$ and edge-set $E(G)$, and we denote $|V(G)|$ and $|E(G)|$ by p and q respectively. Graph Labeling is giving a label at vertices and edges of graph, so that each vertices and each edges having different label. Generally, label uses positive integer.

*This paper has been published on Names Conference 2009, UnLam, Banjarmasin, July 3th - 4st 2009.

Several ways to labeling graph, labeling by domain from set of vertices, set of edges or both which generally referred as vertices labeling, edges labeling and total labeling. At this paper, we will present total labeling, especially edge-magic total labeling (EMTL) on cycle graph.

The edge-magic total labeling is sum of edges label and vertices label which adjacent at edges always equal for each edges. The sum is named magic number with symbol a letter k . Idea of labeling is defined first by [3] at 1960 and here in after [2] made a formula at 1970. A graph is called disconnected if its vertex-set can be partitioned into two subsets, V_1 and V_2 , which have no common element, in such a way that there is no edge with one endpoint in V_1 and the other in V_2 ; if a graph is not disconnected, then it is connected. A cycle graph C_n is connected graph that have n vertex and n edges.

Definition of edge-magic total labeling and super edge-magic total labeling:

Definition 1.1 [4] *Edge-Magic Total Labeling* An edge-magic total labeling of a (p, q) -graph G is bijective function

$$\lambda : V(G) \cup E(G) \rightarrow \{1, 2, 3, \dots, p + q\}$$

such that

$$\lambda(u) + \lambda(uv) + \lambda(v) = k$$

k is a constant for any edge uv of G and is called the magic sum of G . Any graph with an edge-magic total labeling will be called edge-magic. Moreover, λ is a super-edge-magic total labeling of G if $\lambda(V(G)) = \{1, 2, 3, \dots, p + q\}$, and G is said to be super-edge-magic.

For each graph, we number all vertex and edges, we call these numbers positions. Thus, a graph labeling can be represented as a set of ordered pairs of position and its label.

A critical set of a graph G with labeling λ is a set $Q_\lambda = \{(x, y) | x, y \in \{1, 2, 3, \dots, p + q\}\}$, with the ordered pair (x, y) represents label y in position x , which satisfy [5]

- a. λ is the only labeling of G which has label y in position x
- b. No proper subset of Q_λ

2. Edge-Magic Total Labeling

Cycle graph is including graph regular, that is graph having same degree, usually notation of cycle graph with C_n . A cycle graph is connected graph that have n vertices and n edges.

Theorem 2.1 [1] *Every cycle graph C_n has an edge-magic total labeling with range of magic sum is*

$$\frac{5n+3}{2} \leq k \leq \frac{7n+3}{2}, \text{ with } n \text{ is odd}$$

and

$$\frac{5n+4}{2} \leq k \leq \frac{7n+2}{2}, \text{ with } n \text{ is even}$$

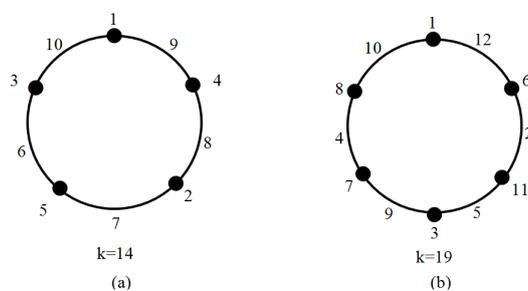


Figure 1: EMTL Cycle graph with $k = 14$ and $k = 19$

Theorem 2.2 *Expanding cycle graph C_n plus a edge has an edge-magic total labeling with range of magic sum is*

$$\frac{5n^2 + 7n + 8}{2(n+1)} \leq k \leq \frac{7n^2 + 17n + 4}{2(n+1)}$$

Proof: Lets S_v denote the sum of vertex labels and S_e denote the sum of the edge labels in a edge-magic labeling λ . In the expanding cycle graph, sum of vertex is $p = n$ and sum of edge is $q = p + 1$. Therefore, k as magic sum of G is

$$\begin{aligned} (n+1)k &= 2S_v + S_e + v_i + v_j \\ &= S_v + 1 + 2 + \dots + (2n+1) + v_i + v_j \\ &= S_v + 2n^2 + 3n + 1 + v_i + v_j \end{aligned}$$

Among the labels, write S for the set $\{v_i | 1 \leq i \leq n\}$ of vertex labels, and k for the sum of elements of S . Then S can consist of the smallest labels, the largest labels, or somewhere in between, so

$$1+2+3+\dots+n+3 \leq S_v + v_i + v_j \leq (n+2) + (n+3) + \dots + (2n+1) + 4n+1$$

or

$$\frac{n^2 + n + 6}{2} \leq S_v + v_i + v_j \leq \frac{3n^2 + 11n + 2}{2}$$

or

$$\frac{5n^2 + 7n + 8}{2(n+1)} \leq k \leq \frac{7n^2 + 17n + 4}{2(n+1)}$$

Theorem 2.3 *Expanding cycle graph C_n plus a edge has an super edge-magic total labeling with range of magic sum is*

$$\frac{5n^2 + 7n + 12}{2(n+1)} \leq k \leq \frac{5n^2 + 11n + 4}{2(n+1)}$$

Proof: Lets S_v denote the sum of vertex labels and S_e denote the sum of the edge labels in a edge-magic labeling λ . In the expanding cycle graph, sum of vertex is $p = n$ and sum of edge is $q = p + 1$. Therefore, k as magic sum of G is

$$\begin{aligned} (n+1)k &= 2S_v + S_e + v_i + v_j \\ &= 2(1+2+\dots+n) + (n+1) + (n+2) + \dots + (2n+1) + v_i + v_j \\ &= 5n^2 + 7n + 6 + v_i + v_j \end{aligned}$$

Among the labels, write S for the set $\{v_i | 1 \leq i \leq n\}$ of vertex labels, and k for the sum of elements of S . Then S can consist of the smallest labels, the largest labels, or somewhere in between, so

$$1+2 \leq v_i + v_j \leq n + (n-1)$$

or

$$3 \leq v_i + v_j \leq 2n-1$$

or

$$\frac{5n^2 + 7n + 12}{2(n+1)} \leq k \leq \frac{5n^2 + 11n + 4}{2(n+1)}$$

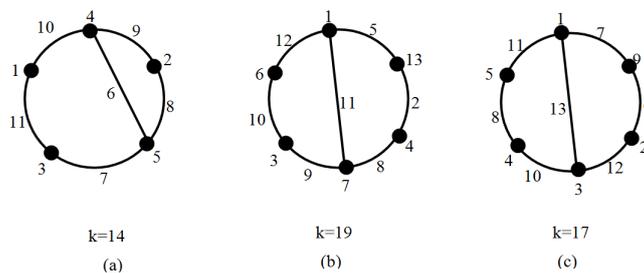


Figure 2: EMTL Expanding Cycle Graph with $k = 14$, $k = 19$ and $k = 17$

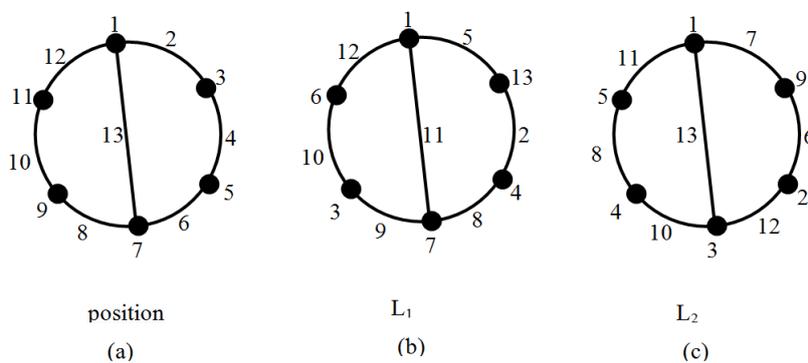


Figure 3: Positon and Labeling of Expanding Cycle Graph

3. Critical set

Figure 3(a), shows of position label of vertices and edges, such that we have edge-magic total labeling of expanding cycle (Figure 3 (b)) with position is

$$L_1 = \{(1, 1), (2, 5), (3, 13), (4, 2), (5, 4), (6, 8), (7, 7), (8, 9), (9, 3), (10, 10), (11, 6), (12, 12), (13, 11)\}$$

and position label of vertices and edges, such that we have edge-magic total labeling with position of another expanding cycle (Figure 3 (b)) is

$$L_1 = \{(1, 1), (2, 7), (3, 9), (4, 6), (5, 2), (6, 12), (7, 3), (8, 10), (9, 4), (10, 6), (11, 5), (12, 11), (13, 13)\}$$

If Q is critical set and minimal number of critical set in expanding cycle graph that have many variation, big more and more expanding cycle more and more many

critical set members and critical set of expanding cycle are

$$Q_1 = \{(1,1), (2,5)\}$$

$$Q_2 = \{(1,1), (2,7)\}$$

obviously after reconstruction, Q_2 have more than one edge-magic total labeling, which are

$$L_{11} = \{(1,1), (2,7), (3,9), (4,6), (5,2), (6,12), (7,3), \\ (8,10), (9,4), (10,6), (11,5), (12,11), (13,13)\}$$

$$L_{12} = \{(1,1), (2,7), (3,12), (4,3), (5,5), (6,9), (7,6), \\ (8,4), (9,10), (10,2), (11,8), (12,11), (13,13)\}$$

$$L_{13} = \{(1,1), (2,7), (3,12), (4,5), (5,3), (6,9), (7,8), \\ (8,2), (9,10), (10,4), (11,6), (12,13), (13,11)\}$$

so Q_2 is not critical set because the reconstruction is not unique, but Q_1 is critical set because reconstruction of Q_1 only one construction.

4. Conclusion

Critical set of expanding cycle graph depended on position and label of position, but expanding cycle graph in form of circle hence distance of position which and label influencing from to the number of members of critical set.

References

- [1] CHAIRUL IMRON, BANDUNG A.S., *Magic Graph on Cycle, The First International Conference on Mathematics and Statistics*, UNISBA, Bandung, 19-21 Juni 2006.
- [2] KOTZIG, A. AND A. ROSA, *Magic Valuations of Finite Graph*, Canad. Math. Bull. 13, 451-461, 1970.
- [3] SEDLACEK, J., problem 27, *Theory of Graphs and it's Applications*, 163-164, Publ. House Czechoslovak Acad. Sci., Prague, 1964
- [4] WALLIS, W.D., E.T. BASKORO, M. MILLER AND SLAMIN, *Edge-Magic Total Labelings*, *Australian Journal of Combinatorics* 22, 177-190, 2000.
- [5] E.T. BASKORO, *Critical Sets in Edge-Magic Total Labelings*, 2005.

INDEKS KEMAMPUAN PROSES BERDASARKAN PROPORSI PERSESUAIAN UNTUK DISTRIBUSI NON NORMAL

Laksmi P Wardhani¹, Resty Z Fahrida, Nur Hasanah

Jurusan Matematika FMIPA ITS Surabaya

¹wmarbun@indo.net.id

Abstrak

Indeks Kemampuan proses adalah suatu alat untuk menganalisa kemampuan proses produksi apakah sesuai dengan spesifikasi yang diberikan. Indeks C_p , C_{pk} , C_{pm} , dan C_{pmk} adalah indeks yang sering digunakan untuk mengukur kemampuan proses dengan berdasarkan proporsi bagian yang tidak sesuai (*Proportion non Conforming*) dan pada indeks ini data diasumsikan berdistribusi Normal.

Indeks Kemampuan proses berdasarkan bagian yang sesuai (*proportion of conforming*) yaitu C_{pc} dapat mengatasi kelemahan-kelemahan dari indeks kemampuan proses yang berdasarkan *Proportion non Conforming* diatas dimana data tidak harus berdistribusi Normal. Dalam tulisan ini dibahas analisis tentang indeks C_{pc} dengan menggunakan asumsi berdistribusi Non Normal yaitu distribusi Poisson dan distribusi Eksponensial berikut estimasi dari C_{pc} yaitu \hat{C}_{pc} .

Katakunci: *Indeks kemampuan proses, distribusi Poisson, distribusi Eksponensial, MLE, limit kepercayaan.*

1. Pendahuluan

Pengendalian Kualitas terhadap suatu produk menjadi hal yang penting dalam dunia industri disebabkan adanya berbagai penyimpangan yang sering terjadi dalam suatu proses produksi. Berbagai penyimpangan terjadi apabila produk yang dihasilkan tidak sesuai dengan dengan spesifikasi yang telah ditetapkan. Setelah proses produksi terken-dali maka dilakukan analisis kemampuan proses, yaitu analisi tentang kemampuan proses untuk memproduksi suatu produk apakah sesuai dengan batas-batas spesifikasi yang telah ditetapkan perusahaan. Analisis kemampuan proses memerlukan indeks kemampuan proses untuk mengukur dan menganalisis potensial dan performa dari suatu proses produksi.

Indeks kemampuan proses sering digunakan adalah C_p , C_{pk} , C_{pm} , dan C_{pmk} dengan dasar perhitungannya berdasarkan *proportion of non conforming* (bagian yang tidak sesuai) . Kelemahan indeks indeks tersebut adalah peubah acak yang digunakan harus ber-distribusi Normal dan batas spesifikasi bawah (BSB) dan batas spesifikasi atas(BSA) harus diketahui terlebih dahulu.

Telah diusulkan sebuah indeks kemampuan proses yang baru untuk mengatasi kelemahan dari indeks-indeks sebelumnya, yaitu indeks C_{pc} . Indeks C_{pc} adalah indeks kemampuan proses yang berdasarkan *proportion of conforming* (bagian yang sesuai). Salah satu kelebihan indeks C_{pc} adalah dapat digunakan untuk data yang berdistribusi non Normal, baik diskrit ataupun kontinu. Pada tulisan ini akan dianalisis indeks C_{pc} untuk asumsi distribusi diskrit yaitu distribusi Poisson dan distribusi kontinu yaitu distribusi Eksponensial. Analisa yang akan dilakukan adalah estimator dari C_{pc} dan limit kepercayaannya.

2. Indeks C_{pc}

Indeks C_{pc} yaitu indeks kemampuan proses berdasarkan *proportion of conforming* (bagian yang sesuai) atau jumlah unit produksi yang memenuhi satu atau beberapa spesifikasi dari proses produksi tersebut/tidak cacat. Indeks C_{pc} didefinisikan sebagai berikut [4]:

$$C_{pc} = \frac{1 - p_0}{1 - p} \quad (1)$$

Dengan p : bagian yang sesuai
 p_0 : proses minimum yang diijinkan.

Nilai p_0 adalah 0,9973, yaitu luasan dari sebaran alamiah dari suatu proses yang besarnya 6σ . Nilai-nilai dibawah atau diatas 0,9973 diijinkan, namun hasil yang diperoleh menjadi tidak maksimum.

Selain itu, indeks C_{pc} juga didefinisikan sebagai berikut:

$$C_{pc} = \min \left\{ \frac{p_0^L}{p_1}, \frac{p_0^U}{p_2} \right\}$$

Dengan $p_1 = P(X < U)$, U batas spesifikasi atas.

$p_2 = P(X > L)$, L batas spesifikasi bawah.

Sedangkan p_0^L dan p_0^U adalah bagian proses yang cacat (tidak sesuai) yang dapat diterima sesuai dengan batas spesifikasi bawah dan atas yang diberikan.

Nilai minimum yang yang diperkenankan dalam proporsi yang sesuai p_0 secara intuisi dikehendaki yang mendekati nilai 1. Walaupun nilai p_0 tidak spesifik, tetapi nilai 0,9973 lebih senang dipilih karena mempunyai sifat-sifat yang mendekati satu dan proses rata-rata tepat berlokasi pada titik tengah spesifikasi $M = (L+U)/2$ dan nilai tersebut merupakan nilai standar dari Indeks Kemampuan Proses [4]. Dalam tulisan ini, supaya memudahkan diasumsikan nilai $p_0 = 0,9973$. Jadi indeks C_{pc} dalam persamaan (1) dapat ditulis kembali dalam bentuk berikut:

$$C_{pc} = \frac{0,0027}{1 - p} \quad (2)$$

Dengan demikian analisa indeks ini selanjutnya dapat dimodifikasi sesuai asumsi nilai p_0 lain yang diinginkan, sedangkan tulisan ini analisa selanjutnya akan memakai persamaan (2).

Pengujian nilai yang berbeda untuk nilai p , akan menghasilkan nilai C_{pc} sebagai berikut:

- (a) Jika nilai $p = 0,9973$, maka nilai indeks $C_{pc} = 1$
- (b) Jika nilai $p > 0,9973$, maka nilai indeks $C_{pc} > 1$
- (c) Jika nilai $p < 0,9973$, maka nilai indeks $C_{pc} < 1$

Satu-satunya parameter yang tidak diketahui dalam indeks C_{pc} adalah p , nilai p tergantung pada bentuk distribusi yang digunakan dalam proses produksi, dan dapat diestimasi dalam bentuk sampel acak dalam proses yang diuji. Hasil estimasinya dinotasikan dalam bentuk \hat{p} , sehingga estimator indeks C_{pc} adalah

$$\hat{C}_{pc} = \frac{0,0027}{1 - \hat{p}} \quad (3)$$

3. Analisa inferensia indeks C_{pc} untuk proses berdistribusi Poisson

Dalam bagian ini indeks C_{pc} dianalisa dalam kasus dimana proses produksi yang diberikan berkarakteristik distribusi Poisson dengan parameter λ . Jika nilai U diketahui, maka indeks dinotasikan dengan C_{pcu} , demikian juga jika nilai L yang diberikan, maka indeks dinotasikan dengan C_{pcl} . Indeks C_{pcu} didefinisikan sebagai [2], [5]:

$$C_{pcu} = \frac{0,0027}{1 - p_1}, \quad \text{dengan } p_1 = P(X < U) \quad (4)$$

Indeks C_{pcl} didefinisikan sebagai

$$C_{pcl} = \frac{0,0027}{1 - p_2}, \quad \text{dengan } p_2 = P(X > L) \quad (5)$$

dengan p_1 dan p_2 merupakan Fungsi Distribusi Kumulatif (CDF) dari distribusi Poisson dengan parameter λ , jadi

$$p_1 = P(X \leq U - 1) = \sum_{x=0}^{U-1} \frac{e^{-\lambda} \lambda^x}{x!} \quad (6)$$

Karena adanya hubungan antara Distribusi Poisson dengan distribusi Gamma maka persamaan (6) dapat ditulis sebagai,

$$p_1 = P(X \leq U - 1) = P(\chi_{2U}^2 > 2\lambda)$$

Dengan χ_{2U}^2 distribusi chi-kuadrat dengan derajat kebebasan $2U$. Oleh karena itu berdasarkan sifat-sifat probabilitas Persamaan (4) dapat ditulis kembali sebagai

$$C_{pcu} = \frac{0,0027}{P(\chi_{2U}^2 < 2\lambda)} \quad (7)$$

Dengan cara yang serupa, indeks C_{pcl} dapat ditulis kembali dalam bentuk

$$C_{pcl} = \frac{0,0027}{P(\chi_{2(L+1)}^2 > 2\lambda)} \quad (8)$$

3.1. Estimasi C_{pcu} dan C_{pcl} untuk proses distribusi Poisson

Jika parameter λ tidak diketahui maka harus diganti dengan estimasinya seperti diketahui MLE (*Maximum Likelihood Estimation*) dari parameter λ untuk n sampel acak (X_1, X_2, \dots, X_n) dari distribusi Poisson adalah mean (rata-rata dari sampel) yaitu $\hat{\lambda} = \bar{X}$. Jadi estimator untuk C_{pcu} dan C_{pcl} masing-masing adalah [2], [5]:

$$C_{pcu} = \frac{0,0027}{P(\chi_{2U}^2 < 2\hat{\lambda})} \quad \text{dan} \quad C_{pcl} = \frac{0,0027}{P(\chi_{2(L+1)}^2 > 2\hat{\lambda})} \quad (9)$$

Untuk menganalisis apakah estimator untuk C_{pcu} dan C_{pcl} yaitu \hat{C}_{pcu} dan \hat{C}_{pcl} merupakan estimator yang tak bias, dilihat terlebih dahulu nilai harapan dari estimator tersebut.

$$E(\hat{C}_{pcu}) = C_{pcu} \times \frac{1}{1 - [e^{-n\lambda}(n\lambda + (n\lambda)^{-2})]}$$

Sehingga estimator tak bias dari C_{pcu} adalah $b_n \times \hat{C}_{pcu}$ dengan b_n sebagai faktor koreksi, dan $b_n = 1 - [e^{-n\lambda}(n\lambda + (n\lambda)^{-2})]$.

3.2. Limit kepercayaan untuk indeks C_{pc}

Selang kepercayaan $100(1 - \alpha)\%$ untuk indeks C_{pc} mempunyai lebar selang sebagai berikut:

$$P(L \leq C_{pc} \leq U) = 1 - \alpha$$

Untuk membangun selang kepercayaan, dilakukan pendekatan dengan distribusi sampling Chi-Kuadrat. Jika diberikan n sampel acak (X_1, X_2, \dots, X_n) dengan $X_i \sim POI(\lambda)$ maka $Y = \sum_{i=1}^n x_i$ juga berdistribusi Poisson dengan parameter $n\lambda$. Karena ada hubungan antara distribusi Poisson dengan distribusi Chi-Kuadrat seperti pada awal tulisan diatas, selang kepercayaan $100(1 - \alpha)\%$ untuk indeks $n\lambda$ adalah

$$\frac{1}{2} \chi_{2Y}^2, \alpha < n\lambda < \frac{1}{2} \chi_{2(Y+1)}^2, 1 - \alpha \quad (10)$$

Dengan cara yang serupa, didapat limit kepercayaan atas $100(1 - \alpha)\%$ untuk $n\lambda$ yaitu

$$P\left(n\lambda < \frac{1}{2}\chi_{2(Y+1),1-\alpha}^2\right) = 1 - \alpha$$

atau

$$P\left(2\lambda < \frac{\chi_{2(Y+1),1-\alpha}^2}{n}\right) = 1 - \alpha \quad (11)$$

Selang kepercayaan $100(1 - \alpha)\%$ untuk C_{pcl} didapat dari persamaan (8) dan (11) didapat adalah

$$\frac{0,0027}{P\left(\chi_{2U}^2 < \frac{\chi_{2(Y+1),1-\alpha}^2}{n}\right)} \quad (12)$$

Dengan cara yang similar didapat selang kepercayaan $100(1-\alpha)\%$ untuk C_{pcu} sebagai berikut:

$$\frac{0,0027}{1 - P\left(\chi_{2(L+1)}^2 < \frac{\chi_{2Y,\alpha}^2}{n}\right)} \quad (13)$$

4. Analisis inferensia indeks C_{pc} untuk proses berdistribusi Eksponensial

Dalam bagian ini indeks C_{pc} dianalisa dalam kasus dimana proses produksi yang diberikan berkarakteristik distribusi Eksponensial dengan parameter θ .

Distribusi Eksponensial diasumsikan sangat beralasan karena dipakai pada uji reliabilitas tentang uji hidup suatu produk . Karena distribusi ini non normal maka akan dilihat selang kepercayaannya satu sisi seperti pada distribusi Poisson pada subbab 3.

Dengan menggunakan batasan seperti persamaan (4) dan (5) didapat masing-masing indeks C_{pcu} dan C_{pcl} pada distribusi Eksponensial adalah [3], [4]:

$$C_{pcu} = \frac{0,0027}{e^{-\theta U}} \quad (14)$$

dan

$$C_{pcl} = \frac{0,0027}{1 - e^{-\theta L}} \quad (15)$$

Dengan CDF untuk $X \sim EXP(\theta)$ adalah $F(x, \theta) = 1 - e^{-\theta x}$, $x > 0$.

4.1. Estimasi C_{pcu} dan C_{pcl} untuk proses distribusi eksponensial

Selanjutnya dicari estimator untuk C_{pcu} dan C_{pcl} dari asumsi distribusi Eksponensial dengan cara mencari estimator untuk parameter θ . M L E untuk θ adalah $\hat{\theta} = \frac{1}{\bar{X}}$, sehingga diperoleh estimator untuk C_{pcu} dan C_{pcl} masing-masing adalah,

$$\hat{C}_{pcu} = \frac{0,0027}{e^{-\hat{\theta}U}} = \frac{0,0027}{e^{-\frac{U}{\bar{X}}}} \quad (16)$$

dan

$$\hat{C}_{pcl} = \frac{0,0027}{1 - e^{-\hat{\theta}L}} = \frac{0,0027}{1 - e^{-\frac{L}{\bar{X}}}} \quad (17)$$

Dengan menggunakan pdf (*probability density functions*) dari distribusi Eksponensial, metode transformasi satu-satu [1] dan $Y = \sum_{i=1}^n X_i \sim GAM(\theta, n)$ didapat distribusi dari \hat{C}_{pcu} dan \hat{C}_{pcl} .
Jika $W = \hat{C}_{cpu}$ maka $W = 0,0027e^{\bar{X}^{-1}U}$ dan pdf dari W yaitu

$$f_W(w) = \frac{(\theta n U)^n}{\Gamma(n)} (\ln w + 5,9145)^{-n-1} w^{-1} e^{(-\theta n U)/(\ln w + 5,9145)},$$

untuk $0,0027 < w < \infty$. Jika $V = \hat{C}_{cpl}$, maka dengan cara yang similar seperti W didapat pdf untuk V adalah [3]:

$$f_V(v) = \frac{(\theta n L)^n}{\Gamma(n)} \left(-\ln \left(\frac{1 - 0,0027}{v}\right)\right)^{-n-1} e^{(\theta n L)/(\ln(1-0,0027/v))} \frac{0,0027}{v^2 - 0,0027v},$$

untuk $0,0027 < w < \infty$.

4.2. Limit Kepercayaan untuk indeks C_{pc}

Selang kepercayaan $100(1 - \alpha)\%$ untuk indeks C_{pc} mempunyai lebar selang sebagai berikut:

$$P(L \leq C_{pc} \leq U) = 1 - \alpha$$

Untuk membangun selang kepercayaan, dilakukan pendekatan dengan distribusi sampling. Jika diberikan n sampel acak (X_1, X_2, \dots, X_n) dengan $X_i \sim EXP(\theta)$ maka $Y = \sum_{i=1}^n X_i \sim GAM(\theta, n)$ dan $Z = \theta Y$ didapat

$Z \sim GAM(1, n)$ terlihat Z tidak bergantung dari parameter θ . Limit kepercayaan bawah $100(1 - \alpha)\%$ untuk θ diperoleh dengan menggunakan percentil α dari distribusi Z yaitu c_1 , sehingga dengan melihat persamaan (14) dan (15) didapat

$$P(Z > c_1) = P(\theta > \frac{c_1}{Y}) = P(\frac{c_1}{Y} < \theta) = P(e^{\frac{-Uc_1}{Y}} < e^{-\theta U}) =$$

$$P\left(C_{pcu} > \frac{0,0027}{e^{\frac{-Uc_1}{Y}}}\right) = 1 - \alpha$$

Jadi limit kepercayaan bawah $100(1 - \alpha)\%$ untuk C_{pcu} adalah

$$\frac{0,0027}{e^{\frac{-Uc_1}{\sum x_i}}}$$

Dengan cara yang similar didapat limit kepercayaan bawah $100(1 - \alpha)\%$ untuk C_{pcl} adalah $\frac{0,0027}{1 - e^{\frac{-Uc_2}{\sum x_i}}}$ dengan c_2 adalah percentile $1 - \alpha$ dari distribusi Z .

5. Kesimpulan

Dari kajian diatas terlihat bahwa indeks kemampuan Proses C_{pc} banyak menanggulangi beberapa kekurangan dari indeks kemampuan Proses yang lama antara lain hanya diasumsikan untuk distribusi normal, sedangkan indeks C_{pc} dapat digunakan untuk distribusi Non-Normal baik untuk proses diskrit seperti distribusi Poisson dan kontinyu seperti distribusi Eksponensial.

Estimasi Parameter dari Indeks C_{pc} untuk kedua distribusi cukup dengan melihat MLE dari parameter distribusi tersebut, sedangkan limit kepercayaan hanya untuk satu sisi saja.

Pustaka

- [1] Bain, L.J. dan Engelhardt, M., *Introduction to Probability and Mathematical Statistics*, 2nd edition, Duxbury Press, California, 1991.

-
- [2] Fahrida, R. Z., *Kajian Indeks Kemampuan Proses C_{pc} dengan menggunakan asumsi distribusi Poisson*, Tugas Akhir, Jurusan Matematika FMIPA ITS, 2008.
- [3] Nur, H., *Kajian Indeks Kemampuan Process C_{pc} dengan menggunakan asumsi distribusi Eksponensial* Tugas Akhir. Jurusan Matematika FMIPA ITS, 2009.
- [4] Perakis, M. dan Xekalaki, E. 2002. *A Process Capability Index that it based on the Proportion of Conformance*. Journal Statistical Computation and Simulation, 2002, Vol 72(9), pp. 707-718.
- [5] Perakis, M. dan , Xekalaki, E., *A Process Capability Indeks For Discrete Processes*. Journal Statistical Computation and Simulation, March 2005, vol 75 No.3, pp 175-187.

OPTIMASI DALAM PENENTUAN DOSIS OPTIMAL PADA KEMOTERAPI TUMOR

Yopi Andry Lesnussa¹, Subchan²

¹Jurusan Matematika, FMIPA Universitas Pattimura Ambon

²Jurusan Matematika, FMIPA ITS Surabaya

¹yopi_lesnussa@yahoo.com, ²subchan@matematika.its.ac.id

Abstrak

Konstruksi model matematis dari suatu fenomena dalam bidang matematika biologi merupakan hal yang sangat penting. Salah satunya dapat diterapkan dalam proses kemoterapi tumor. Sebagai salah satu penyakit yang mematikan, pengobatan kemoterapi tumor perlu dioptimalkan untuk mencegah proliferasi sel yang tidak terkontrol. Namun proses kemoterapi yang tidak tepat, dapat berakibat fatal bagi pasien penyakit tumor. Sehingga interval waktu dan dosis yang tepat dalam kemoterapi sangat efektif untuk mengurangi ukuran tumor. Permasalahan kemoterapi tumor dimodelkan sebagai permasalahan optimal dimana penentuan dosis obat optimum merupakan fungsi tujuannya. Permasalahan optimal selanjutnya ditransformasikan menjadi permasalahan pemrograman nonlinier (PNL), yang selanjutnya diselesaikan dengan software pemrograman nonlinier (PNL).

Katakunci: *Kendali optimum, kemoterapi tumor, dosis optimum.*

1. Pendahuluan

Tumor adalah segolongan penyakit paling berbahaya yang ditandai dengan pembelahan sel yang tidak terkendali. Sel-sel tersebut mampu menyerang jaringan biologis lainnya dengan pertumbuhan langsung di jaringan yang bersebelahan (invasi) atau migrasi sel ke tempat yang jauh (metastasis). Pertumbuhan yang tidak terkendali tersebut menyebabkan mutasi di gen vital yang mengendalikan pembelahan sel. Beberapa mutasi dapat mengubah sel normal menjadi sel tumor. Penyakit tumor disebabkan oleh beberapa faktor, antara lain : virus, kecanduan rokok, radiasi sinar ultraviolet, zat kimia, makanan berlemak, faktor keturunan, dan lain-lain [5], [1]. Penyakit tumor merupakan salah satu penyebab kematian terbesar di dunia, oleh karena itu perlu perawatan sejak awal. Beberapa jenis pengobatan penyakit tumor antara lain: bedah (operasi), radioterapi, kemoterapi, terapi hormon, imunoterapi, dan kombinasinya [8]. Kemoterapi adalah tindakan/terapi pemberian senyawa kimia (obat tumor) untuk mengurangi, menghilangkan atau menghambat pertumbuhan parasit atau mikroba di tubuh pasien (hospes). Obat-obatan yang sering digunakan dalam kemoterapi misalnya golongan siklofosfamid, methotreksat, dan beberapa obat sitotoksik seperti amsacrine, cisplatin, cyclophosphamide, dan lain-lain. Kemoterapi dengan dosis obat yang berlebihan atau tidak tepat dapat membunuh atau merusak jaringan dan sel tubuh yang normal, serta menyebabkan efek samping bagi penderita penyakit tumor seperti lemas, mual dan muntah, gangguan pencernaan, rambut rontok, dan lain-lain. Obat tumor merupakan obat spesialisik sehingga hanya dibenarkan penggunaannya oleh dokter yang berpengalaman di bidang pengobatan ini. Karena itu durasi dan dosis obat yang tepat sangat penting dalam pengobatan tumor [3].

Di bidang matematika biologi, fenomena ini dapat diselesaikan dengan mengkonstruksi suatu model matematis dan menerapkan teori kendali optimal untuk menentukan kapan dan sejauh mana dosis yang tepat dalam proses kemoterapi tumor [9], [4]. Beberapa penelitian di bidang kendali optimum untuk kemoterapi tumor antara lain : [11] meneliti tentang interaksi teori kendali optimum dengan kemoterapi tumor yang meliputi 3 bidang yaitu melibatkan model kinetik pertumbuhan miscellaneous, model siklus sel dan klasifikasi model. [6] meneliti obat-obatan anti tumor yang ditujukan untuk meminimalkan ukuran tumor dimana secara anali-

tik gradien dari semua konstrain dikonstruksikan dan masalah diselesaikan secara umum dengan menggunakan persamaan Gompertz dan persamaan Bellman. [7] meneliti tentang teori kendali optimum untuk menganalisis bagaimana menghitung pengaruh negatif dan kendala dari tumor pada sel normal yang mempengaruhi penerapan obat optimum dalam kemoterapi tumor dan menentukan aturan optimum yang meminimalkan sel tumor pada akhir periode terapi dengan mempertahankan populasi sel normal. [2] meneliti tentang hubungan kendali optimal dengan terapi obat dan menguji atau membandingkan berbagai strategi pengendalian optimal termasuk kendali kuadrat, kendali linier dan ruang kendala. [8] meneliti tentang masalah kendali optimum yang dirumuskan dan diselesaikan untuk model sel cycle nonspesifik dan sel cycle spesifik sehingga mendapatkan jadwal kemoterapi yang efektif untuk meminimalkan ukuran tumor dan membatasi kerusakan pada sel normal. Simulasi model pertumbuhan tumor dilakukan untuk mengetahui pola pertumbuhan tumor dan aplikasinya untuk meningkatkan terapi tumor supaya dapat untuk memahami dinamika respon obat dalam tubuh. Kendali optimal kemoterapi sangat dibutuhkan untuk mengoptimalkan efek pemberian obat dengan cara mengatur dosis obat dan rentang waktu pemberian obat, sehingga dalam penelitian ini ditunjukkan untuk bagaimana menentukan interval waktu dan dosis optimal dalam kemoterapi tumor.

2. Hasil dan Pembahasan

2.1. Model matematika kemoterapi tumor

Model matematika yang dikembangkan dalam penelitian ini terdiri dari : tiga model populasi sel, diantaranya sel tumor, sel *effektor-immun* dan sel sirkulasi limposit serta satu model konsentrasi obat dalam peredaran darah. Komponen sistem kekebalan tubuh (*immun*), meliputi populasi sel effektor-immun yang secara aktif berfungsi membunuh sel tumor sedangkan populasi sel sirkulasi limposit berfungsi mengawasi atau menekan peningkatan kerusakan sel akibat efek samping kemoterapi. Model sirkulasi sel limposit mewakili ukuran kesehatan pasien sehingga digunakan sebagai pengganti populasi sel normal. Sistem dari persamaan differensial yang digunakan menggambarkan pertumbuhan, kematian, dan interaksi dari masing-masing

populasi dengan pengobatan kemoterapi yang diberikan oleh sistem berikut [2] :

$$\dot{T} = aT(1 - bT) - c_1NT - K_TMT \quad (1)$$

$$\dot{N} = \alpha_1 - fN + g\frac{T}{h+T}N - pNT - K_NMN \quad (2)$$

$$\dot{C} = \alpha_2 - \beta C - K_CMC \quad (3)$$

$$\dot{M} = \gamma M + V_M(t) \quad (4)$$

Dengan populasi sel tumor, populasi sel kekebalan dan konsentrasi obat dalam waktu t dinotasikan sebagai berikut :

- $T(t)$: Populasi sel tumor
- $N(t)$: Populasi sel *effektor-immun*
- $C(t)$: Populasi sirkulasi limposit
- $M(t)$: Konsentrasi kemoterapi obat

Setiap persamaan diatas memiliki kondisi awal secara umum, yaitu :

$$T(0) = T_0 \geq 0, N(0) = N_0 \geq 0, C(0) = C_0 \geq 0, \text{ dan } M(0) = M_0 \geq 0.$$

Pada persamaan (1) populasi sel tumor diasumsikan tumbuh secara logistik, walaupun sel tumor dibunuh oleh sel *effektor* melalui aksi masa dinamik. Persamaan (2) pada sel *effektor* memiliki laju konstanta sumber α_1 dan $-fN$ merupakan laju kematian seimbang. Sel *effektor* juga direkrut oleh sel tumor melalui persamaan *Michael-Menten* $g\frac{T}{h+T}$. Sirkulasi sel limposit pada persamaan (3) memiliki laju konstanta sumber α_2 dan $-\beta C$ merupakan bentuk kematian seimbang. Pada persamaan (4) $V_M(t)$ menotasikan dosis obat pada konsentrasi kemoterapi dan $-\gamma M$ merupakan kemunduran sistem secara seimbang untuk konsentrasi.

2.2. Analisis sistem dinamik

Analisis Sistem Dinamik Sistem dinamik dari persamaan (1)-(4) dapat dianggap sebagai suatu titik tetap dan dapat ditentukan karakteristik stabilitasnya. Jika $V_M(t)$ adalah suatu konstanta dengan nilai $V_M(t) = V_M$, maka persamaan (4) dapat menjadi :

$$M = \frac{V_M}{\gamma} \quad (5)$$

Tabel 1: Estimasi nilai parameter

Parameter	Unit	Deskripsi	Nilai Estimasi
a	hari ⁻¹	Laju pertumbuhan tumor	$4,31 \times 10^{-3}$
b	sel ⁻¹	$\frac{1}{b}$ adalah kapasitas tumor	$1,02 \times 10^{-14}$
c_1	sel ⁻¹ hari ⁻¹	Bagian sel tumor yang dibunuh oleh sel <i>effektor</i>	$3,41 \times 10^{-10}$
f	hari ⁻¹	Laju kematian sel <i>effektor</i>	$4,12 \times 10^{-2}$
g	hari ⁻¹	Laju rekrutment sel <i>effektor</i> maksimum	$1,50 \times 10^{-2}$
h	sel ²	Koefisien steepnes dari kurva rekrutment sel <i>effektor</i>	$2,02 \times 10^1$
K_C, K_N	hari ⁻¹	Bagian sel <i>effektor</i> dan sirkulasi limposit yang dibunuh oleh kemoterapi	$6,00 \times 10^{-1}$
K_T	sel ⁻¹	Bagian sel tumor yang dibunuh oleh kemoterapi	$8,00 \times 10^{-1}$
p	hari ⁻¹	Laju inaktivasi sel <i>effektor</i> oleh sel tumor	$2,00 \times 10^{-11}$
α_1	sel ⁻¹ hari ⁻¹	Konstanta sumber dari sel <i>effektor</i>	$1,20 \times 10^4$
α_2	sel ⁻¹ hari ⁻¹	Konstanta sumber dari sirkulasi limposit	$7,50 \times 10^8$
β	hari ⁻¹	Laju kematian dari sirkulasi limposit	$1,2 \times 10^{-2}$
γ	hari ⁻¹	Laju penurunan kemoterapi obat	$9,00 \times 10^{-1}$

Substitusi nilai M pada persamaan (5) ke persamaan (3) diperoleh :

$$C = \frac{\alpha_2 \gamma}{\gamma \beta + K_C V_M} \quad (6)$$

Juga untuk persamaan (1) didapat

$$0 = T[a(1 - bT) - c_1N - K_TM] \quad (7)$$

Untuk $T = 0$ terdapat suatu titik keseimbangan (*equilibrium*), sehingga jika disubstitusikan ke persamaan (2), diperoleh :

$$N = \frac{\alpha_1\gamma}{\gamma f + K_NV_M} \quad (8)$$

Untuk analisa kestabilan dari titik keseimbangan $T = 0$, dapat dibentuk matriks Jacobian dari persamaan (1)-(4), sebagai berikut :

$$J = \begin{pmatrix} \chi & -c_1T & 0 & -K_TT \\ -pN + gN\frac{h}{(h+T)^2} & \zeta & 0 & -K_NN \\ 0 & 0 & -\beta - K_CM & -K_CC \\ 0 & 0 & 0 & -\gamma \end{pmatrix}$$

dengan $\chi = -2abT + a - c_1N - K_TM, \zeta = -f - pT - K_NM + g\frac{T}{(h+T)}$.

Dengan mensubstitusi nilai $T = 0$, dapat diperoleh matriks berikut :

$$J = \begin{pmatrix} a - c_1N - K_TM & 0 & 0 & 0 \\ -pN & -f - K_NM & 0 & -K_NN \\ 0 & 0 & -\beta - K_CM & -K_CC \\ 0 & 0 & 0 & -\gamma \end{pmatrix}.$$

Matriks Jacobian diatas dipartisi ke dalam bentuk matriks ordo 2×2 , sebagai berikut :

$$J_{11} = \begin{pmatrix} a - c_1N - K_TM & 0 \\ -pN & -f - K_NM \end{pmatrix} \quad J_{12} = \begin{pmatrix} 0 & 0 \\ 0 & -K_NN \end{pmatrix}$$

$$J_{21} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad J_{22} = \begin{pmatrix} -\beta - K_CM & -K_CC \\ 0 & -\gamma \end{pmatrix}.$$

Setiap matriks hasil pertisi diatas diselesaikan dengan menggunakan

persamaan $J - eI = 0$, sehingga diperoleh nilai eigen sebagai berikut :

$$\begin{aligned} e_1 &= a - c_1 N - K_T M \\ e_2 &= -f - K_N M \\ e_3 &= -\beta - K_C M \\ e_4 &= -\gamma \end{aligned}$$

Substitusi nilai M , C dan N pada persamaan (5) - (8) pada nilai-nilai eigen diatas, diperoleh :

$$\begin{aligned} e_1 &= a - \frac{c_1 \alpha_1 \gamma}{\gamma f + K_N V_M} - \frac{K_T V_M}{\gamma} \\ e_2 &= -f - \frac{K_N V_M}{\gamma} \\ e_3 &= -\beta - \frac{K_C V_m}{\gamma} \\ e_4 &= -\gamma \end{aligned}$$

dimana semua parameter positif e_2, e_3, e_4 bernilai negatif. Sehingga pada titik keseimbangan $T = 0$ berlaku stabil asimtotik lokal, dengan

$$a - \frac{c_1 \alpha_1 \gamma}{\gamma f + K_N V_M} - \frac{K_T V_M}{\gamma} < 0 \quad (9)$$

Ketika $V_M = 0$ (tidak ada pengobatan), substitusi nilai $V_M = 0$ ke M , C , dan N persamaan (5) - (8), maka diperoleh :

$$T = 0, N = \frac{\alpha_1}{f}, C = \frac{\alpha_2}{\beta} \text{ dan } M = 0$$

Titik-titik ini dengan menggunakan nilai-nilai estimasi parameter yang diberikan pada tabel 1, hasilnya tidak akan memenuhi persamaan (9). Tetapi ketika $V_M = 1$ maka diperoleh :

$$T = 0, M = \frac{1}{\gamma}, N = \frac{\alpha_1}{f + \frac{K_N}{\gamma}} \text{ dan } C = \frac{\alpha_2}{\beta + \frac{K_C}{\gamma}}$$

Titik-titik T , M , N dan C , yang baru ini merupakan titik tetap stabil karena memenuhi persamaan (9). Ini mengindikasikan bahwa jika pengobatan diterapkan secara optimal pada waktu yang tepat maka keseimbangan tumor akan semakin stabil atau menuju ke titik keseimbangan nol ($T = 0$).

2.3. Masalah kendali optimum

Masalah kendali optimum yang akan dioptimalkan adalah dosis obat optimal yang bertujuan untuk meminimalkan ukuran populasi sel tumor. Bentuk kendali optimum yang akan diteliti dalam penelitian ini adalah bentuk kendali kuadratik, yang diterapkan pada suatu fungsi objektif untuk meminimalkan ukuran populasi sel tumor, sehingga dapat diperoleh dosis optimal yang tepat. Fungsi objektif (indeks performa) yang akan diminimalkan yaitu :

$$J(V_M) = \int_0^{t_f} \left(T(t) + \frac{\varepsilon}{2} V_M^2(t) \right) dt \quad (10)$$

2.4. Simulasi numerik dan analisa hasil simulasi

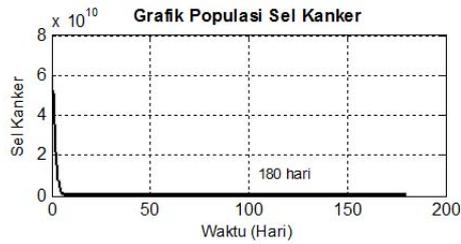
Simulasi dari model matematika untuk populasi sel tumor, sel *effektor-immun*, sel limposit dan konsentrasi obat, diselesaikan dengan menggunakan program DOTcvpSB versi R2010_E3 (Dynamic Optimization Trajectory Control Vector Parameterization System Biology). Hasil simulasi ini bertujuan untuk mendapatkan nilai optimasi secara numerik dari fungsi kendali kuadratik yang mengindikasikan dosis obat optimal, dengan menggunakan estimasi nilai parameter pada Tabel 1.

i. Simulasi secara numerik

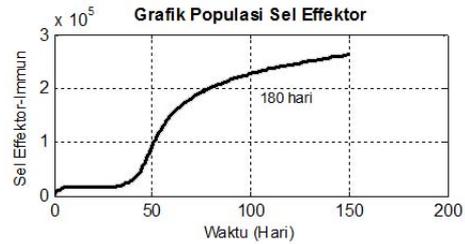
Proses simulasi sel tumor, sel kekebalan dan konsentrasi obat dilakukan dengan waktu awal $t_0 = 0$ dan waktu akhir $t_f = 180$ atau proses simulasi dilakukan selama 180 hari. nilai variabel kendali untuk dosis obat V_M berkisar antara $0 < V_M < 1$. Analisa proses simulasi dibagi dalam 4 kasus, sebagai berikut:

a. Kasus $T_0 > N_0$ dan $T_0 > C_0$.

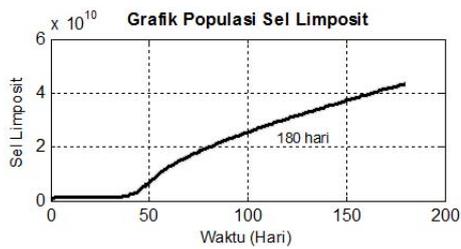
Untuk kondisi awal populasi sel tumor 6×10^{10} sel, populasi sel *effektor-immun* $3,5 \times 10^2$ sel, populasi awal sel limposit $6,25 \times 10^3$ sel. Dari Gambar (1)-(4) dapat dilihat bahwa pada saat populasi sel tumor turun menuju ke titik keseimbangan tumor nol pada awal pengobatan akibat pengaruh obat, sel effektor dan limposit akan tumbuh secara perlahan, sedangkan konsentrasi obat akan



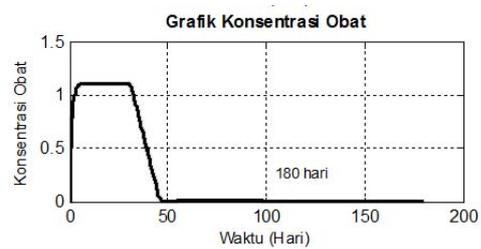
Gambar 1: Populasi sel tumor ($T_0 > N_0$ & $T_0 > C_0$)



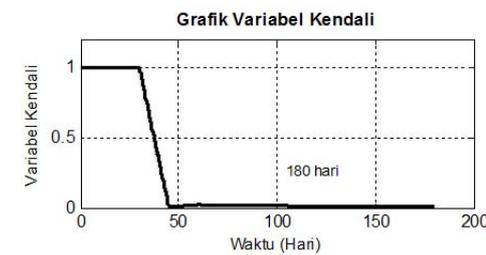
Gambar 2: Populasi sel *effektor* ($T_0 > N_0$ & $T_0 > C_0$)



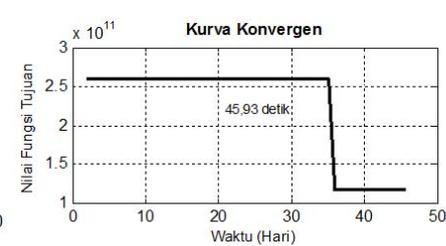
Gambar 3: Populasi sel limposit



Gambar 4: Konsentrasi Obat ($T_0 > N_0$ & $T_0 > C_0$)



Gambar 5: Variabel kendali ($T_0 > N_0$ & $T_0 > C_0$)



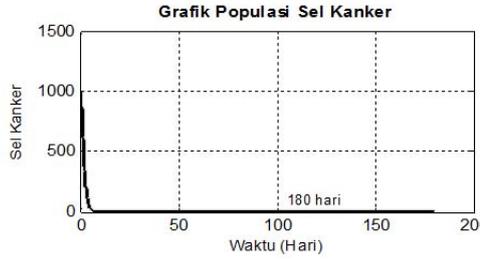
Gambar 6: Kurva konvergen ($T_0 > N_0$ & $T_0 > C_0$)

dikurangi seiring dengan berkurangnya populasi sel tumor. Sedangkan kurva konvergen menunjukkan bahwa fungsi tujuan akan semakin lama konvergen ke titik keseimbangan.

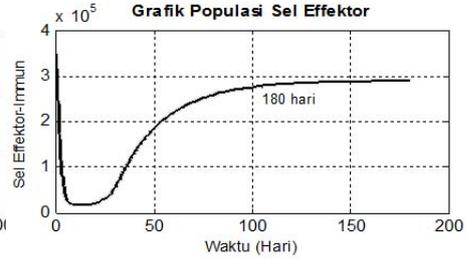
b. Kasus $T_0 < N_0$ dan $T_0 < C_0$.

Untuk kondisi awal populasi sel tumor 10^3 sel, populasi sel *effektor*-

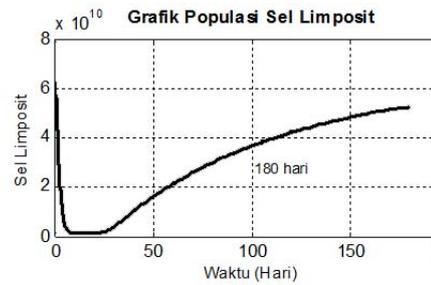
immun $3,5 \times 10^5$ sel, populasi awal sirkulasi limposit $6,25 \times 10^3$ sel.



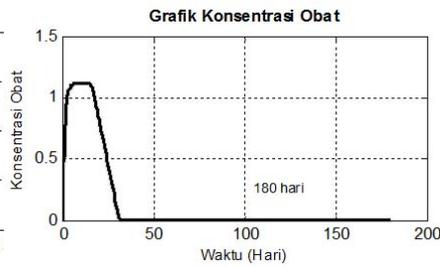
Gambar 7: Populasi sel tumor
($T_0 < N_0$ & $T_0 < C_0$)



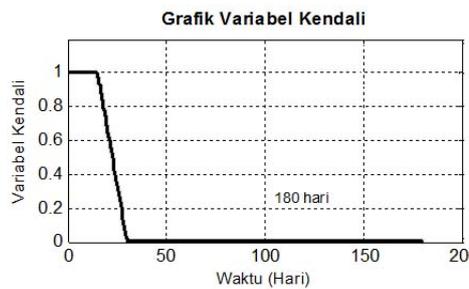
Gambar 8: Populasi sel *effektor*
($T_0 < N_0$ & $T_0 < C_0$)



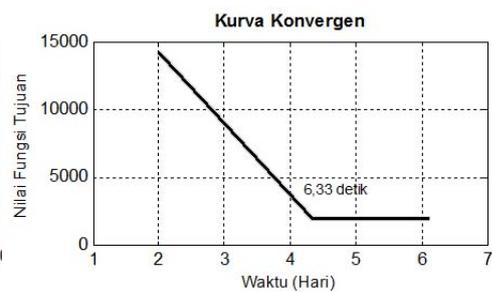
Gambar 9: Populasi sel limposit
($T_0 < N_0$ & $T_0 < C_0$)



Gambar 10: Konsentrasi Obat
($T_0 < N_0$ & $T_0 < C_0$)



Gambar 11: Variabel kendali
($T_0 < N_0$ & $T_0 < C_0$)



Gambar 12: Kurva konvergen
($T_0 < N_0$ & $T_0 < C_0$)

Dari Gambar (7)-(12) dapat dilihat bahwa pada saat populasi sel tumor turun menuju ke titik keseimbangan tumor nol, sel effektor dan limposit akan berkurang juga sampai mendekati titik minimum dan tumbuh secara perlahan, untuk konsentrasi obat akan dikurangi seiring dengan berkurangnya populasi sel tumor. Sedangkan kurva konvergen menunjukkan bahwa fungsi tujuan semakin cepat konvergen ke titik keseimbangan.

Tabel 2: Hasil simulasi numerik untuk dosis obat

No.	Simulasi untuk kasus	Dosis optimum $J(V_M)$
1	$T_0 > N_0$ dan $T_0 > C_0$	116229226840,05139
2	$T_0 < N_0$ dan $T_0 < C_0$	1946,92209964

ii. Analisa hasil simulasi

Dari hasil simulasi secara numerik dengan menggunakan DOTcvpSB diperoleh nilai numerik dari fungsi objektif yang diminimumkan terhadap ukuran populasi sel tumor untuk mendapatkan nilai dosis obat optimal sebagai variabel kendali. Tabel 2, menunjukkan bahwa pada saat populasi sel tumor jauh lebih kecil dari populasi sel kekebalan tubuh maka dosis obat yang dibutuhkan dalam proses kemoterapi lebih kecil dibandingkan dengan dosis obat yang diterapkan untuk populasi tumor yang lebih besar dari populasi sel kekebalan tubuh.

Ini merupakan dosis optimal yang dicapai dalam menekan ukuran populasi sel tumor seminimal mungkin. Proses ini menunjukkan bahwa penggunaan dosis obat dalam suatu periode pengobatan kemoterapi tumor akan dikurangi dosisnya seiring dengan berkurangnya ukuran populasi sel tumor dan meningkatnya populasi sel kekebalan tubuh. Selain itu, peranan sel *effektor-immun* dengan populasinya yang semakin besar juga sangat berpengaruh dalam membunuh dan menekan populasi sel tumor seminimal mungkin. Oleh karena itu, obat yang dipakai dalam proses kemoterapi, selain berfungsi membunuh dan menekan populasi sel tumor diharapkan dapat merangsang pertumbuhan populasi sel kekebalan tubuh. Sehingga pada saat populasi sel tumor mencapai titik keseimbangan nol dan reaksi dari proses pengobatan

kemoterapi berhenti, maka fungsi pertahanan dan kekebalan tubuh dapat digantikan oleh sel kekebalan tubuh.

3. Kesimpulan

Kondisi awal yang mewakili ukuran populasi sel tumor (T_0), populasi *effektor-immun* (N_0) dan populasi sel sirkulasi limposit (C_0) sangat berpengaruh terhadap dosis obat optimal (V_M) yang diterapkan dalam proses kemoterapi tumor. Semakin besar ukuran populasi sel tumor, maka semakin besar pula dosis obat optimal yang diterapkan. Titik keseimbangan (equilibrium) $T = 0$, merupakan titik stabil dan dapat dipenuhi ketika nilai variabel kendali $V_M = 1$ (pengobatan mencapai dosis optimal). Trajektori konsentrasi obat menurun drastis pada saat ukuran populasi tumor mencapai keseimbangan tumor nol, pada saat titik keseimbangan tumor nol maka populasi sel *effektor-immun* dan populasi sel sirkulasi limposit akan meningkat drastis.

Pustaka

- [1] AFENYA E., *Mathematical Model of Cancer and Their Relevant Insights*, Mathematical Biology and Medicine 9, 173-223, 1996.
- [2] DE PHILLIS L.G., GU W., FISTER K.R, HEAD T., MAPLES K., MURUGAN A., NEAL T., DAN YOSHIDA K., *Chemoterapy for Tumors : an Analysis of the Dynamics and a Study of Quadratic and Linear Optimal Control*, Mathematical Biosciences 29, 292-315, 2007.
- [3] HAROLD J.M., DAN PARKER R.S., *Clinically Relevant Cancer Chemoterapy Dose Scheduling via Mixed Integer Optimization*, Computer and Chemical Engineering 33, 2042-2054, 2009.
- [4] ITIK M., SALAMCI M.U. DAN BANKS, S.P., *Optimal Control of Drug Therapy in Cancer Treatment*, Nonlinear Analysis 71, e1473-e1486, 2009.

-
- [5] MACDONALD, F., FORD, C.H.J, DAN CASSON, A.G., *Molecular Biology of Cancer*, Second Edition, Garland Science/BIOS Scientific Publishers, London, 2005
- [6] MARTIN, R.B., *Optimal Control Drug Scheduling of Cancer Chemotherapy*, Pergamon Press Ltd, Automatica 28, 1113-1123, 1992
- [7] MATVEEV A.S., DAN SAVKIN A.V., *Application of Optimal Control Theory to Analysis of Cancer Chemotherapy Regimens*, Systems & Control Letters 46, 311-321, 2002
- [8] PINKY D., VIVEK D., DAN PISTIKOPOULOS, E.N., *Optimal Delivery of Chemotherapeutic Agents in Cancer*, Computers and Chemical Engineering 32, 99-107, 2008
- [9] PREZIOSI, L., *Cancer Modeling and Simulation*, Chapman & Hall/CRC Mathematical Biology and Medicine, New York, 2003.
- [10] SUBCHAN, S., DAN ZBIKOWSKI, R., *Computational Optimal Control Tools and Practise*, John Willey and Sons, Ltd, publication, United Kingdom, 2003.
- [11] SWAN, G.W., *Role of Optimal Control Theory in Cancer Chemotherapy*, Mathematical Biosciences 101, 237-284, 1990.