Pemilihan Model Menggunakan Struktur Perkalian Distribusi

Nur Iriawan
Statistika ITS

ABSTRAK

Pemilihan model terbaik dalam pemodelan suatu data statistik sering terhambat karena tidak terpenuhinya asumsi-asumsi yang diperlukannya. Adanya dua model atau lebih yang tidak tersarang (nested), misalnya, sering juga merupakan kendala dalam pemilihan-nya. Dalam makalah ini dicobakan untuk menyusun model-model yang dibandingkan, untuk dipilih mana yang terbaik, kedalam sebuah struktur perkalian dan selanjutnya dibobotkan sebagai suatu distribusi dengan menambahkan satu parameter dominator. Struktur pemilihan model ini dikembangkan berdasarkan makalah yang ditulis Cox (1962) dengan penerapannya pada Markov Chain Monte Carlo (MCMC). Beberapa contoh yang diberikan menunjuk-kan bahwa struktur ini memberikan cara pemilihan model yang lebih handal.

ABSTRACT

Model selection in the statistical modeling analysis frequently meet some difficulties due to the failure to fulfill the assumption required. Model determination of two or more nested models, for example, is the real problem in the statistical modeling that has to be faced. This paper describes this kind of modeling to compare and choose the best model to represent the given data by constructing a joint model with a dominator, λ, in a multiplicative structure. This structure is developed based on the Cox's (1962) paper and is implemented by using Markov Chain Monte Carlo (MCMC) method. Some illustrations are given here to demonstrate the work of this model selection method.

Kata kunci :Markov Chain Monte Carlo, Marginalisasi, Bayes factor, Multiplicative model, pemilihan model, model tidak tersarang.

1. Pendahuluan

Telah banyak metode-metode statistik yang dikembangkan oleh para peneliti, khususnya dalam kaitannya dengan analisis model linear. Namun masih banyak penelitian-penelitian tersebut yang diarahkan pada cara konvensional, dimana observasi kesalahan pemodelan dianggap mempunyai sifat IIDN$(0,\sigma^2)$. Pendekatan pemodelan seperti ini dapat dilihat pada hasil penelitian Zeger dan Karim (1991), Gelfand, Hills, Racine-Poon dan Smith (1990), Carlin dan Chib (1995), Gilks, Clayton, Spiegelhalter, Best, McNeil, Sharples dan Kirby(1993) serta Wakefield, Smith, Racine-Poon dan Gelfand (1994).

Pada beberapa tahun terakhir telah banyak penelitian yang menekankan pada pengembangan metode pengambilan keputusan dengan menggunakan pendekatan Bayesian dengan memperluas bentuk-bentuk asumsi kesalahan yang cenderung tidak Normal. Pengembangan semacam ini terutama didasari oleh adanya pemikiran tentang kenyataan bahwa model suatu kesalahan pemodelan itu tidak dapat selalu dianggap berbentuk distribusi yang simetris atau Normal, meskipun jumlahan kesalahan-kesalahan tersebut adalah nol. Di banyak kasus bahkan, model kesalahan ini sering dijumpai agak menceng dan tidak mempunyai kecenderungan Normal sama sekali. Box and Tiao (1973) menyederhanakan asumsi ini dengan menggunakan distribusi simetris, Exponential Power distribution, sebagai model distribusi kesalahan dengan parameter penyimpangan terhadap Normalitasnya β. Asumsi distribusi ini akan menganggap bahwa suatu data kesalahan model berdistribusi Normal jika parameter β bernilai nol dan akan merupakan kelas distribusi simetris jika β tidak nol.

Dalam pemilihan dan pembandingan beberapa model, Kass and Raftery (1995), O'Ha-gan (1995), and Carlin and Chib (1995) telah menemukan metode yang sangat terkenal dan sering dirujuk para peneliti lain. Disini penulis akan menggunakan metode mereka sebagai dasar pengembangan untuk menciptakan metode pemilihan model baru dengan membentuk distribusi gabungan dari beberapa model dengan menggunakan asas perkalian yang tidak mensyaratkan asumsi Normalitas pada setiap error-nya. Distribusi ini kemudian akan dinamakan sebagai Struktur Perkalian Distribusi gabungan (SPD). Ide distribusi gabungan ini pertama kali dilontarkan oleh Cox (1962) dan Carota, Parmigiani dan Polson (1996). Meskipun SPD ini akan tampak sangat kompleks dan sulit untuk diselesaikan secara analitis, namun secara teknis, metode SPD akan mudah dihitung dengan menggunakan estimasi Bayes factor yang menggunakan pendekatan Markov Chain Monte Carlo (MCMC) dari pada cara konvensional non-Bayesian ((Geman and Geman,1984),(Casella and George, 1992)).

Untuk menunjukkan hasil dari penelitian ini, penulis pertama kali akan menggunakan SPD dengan MCMC ini dalam pemilihan model distribusi suatu data $\boldsymbol{Y}$, yang akan diadukan antara Eksponensial dan Log-Normal. Berikutnya, metode ini akan ditunjukkan kesahihannya dalam pemilihan model regresi linear tidak bersarang.

2. Struktur Perkalian Distribusi (SPD)

Di beberapa tahun belakangan ini telah banyak peneliti yang mengembangkan metode pemilihan model dengan pendekatan Bayesian. Ukuran pemilihannya didasarkan pada nilai Bayes faktor yang diperoleh. Penggunaan *harmonic mean* dari sampel suatu probabilitas bersyarat dari posterior likelihoodnya dapat dilihat pada makalah yang ditulis oleh Newton dan Raftery (1994). Carlin, Polson dan Stopper (1992) serta Carlin dan Chib (1995) menggunakan indikator model M sebagai sebuah parameter dalam pemilihan modelnya dalam proses samplingnya. Pendekatan lain dapat dilihat pada Gelfand dan Dey (1994), Kass dan Raftery (1995) serta dalam Chib (1995). Termotivasikan oleh pengembangan metode pemilihan model dengan cara membangkitan model terpilih, M_g, melalui implementasi MCMC pada probabilitas bersyarat posterior likelihood untuk mengestimasi nilai Bayes faktor seperti dalam Carlin dan Chib (1995), model SPD ini dikembangkan.

2.1 Pengembangan model statistik SPD

Anggap ada suatu data yang tampak berasal dari dua macam distribusi atau model yang berbeda, misalkan $f_1(x,\theta)$ dan $f_2(x,\omega)$. Keputusan dari distribusi atau model mana yang lebih tepat dan lebih representatif untuk memodelkan data tersebut adalah sangat penting sebelum analisis dilanjutkan. Disini alat uji tradisional yang telah ada; seperti : pemilihan model dengan dasar analisis pada jenis data Normal atau $f_1(x,\theta)$ dan $f_2(x,\omega)$ harus merupakan dua model yang tersarang (nested); sudah kurang tepat untuk digunakan. Untuk mengatasi hal ini, perkalian dua model setelah masing-masing dipangkatkan dengan suatu parameter λ sebagai sebuah density baru dikembangkan.

Dengan menganggap bahwa kedua model pendekatan di atas, $f_1(x,\theta)$ dan $f_2(x,\omega)$, adalah independen, density yang baru ini dinamakan sebagai model SPD atau $f_{SPD}(x,\lambda,\theta,\omega)$. Selanjutnya, parameter λ digunakan sebagai indikator dominasi model dalam $f_{SPD}(x, \lambda, \theta, \omega)$. Metode ini pertama kali dikembangkan oleh Cox (1962) dan pernah dibahas dalam Carota et al. (1996). Model dasarnya dapat dituliskan sebagai berikut :

$$f_{SPD}\left(x,\lambda,\theta,\omega\right) = C\left(\lambda,\theta,\omega\right)\, f_1^{\lambda}(x,\theta)\, f_2^{1-\lambda}(x,\omega) \quad (1)$$

Dimana $C(\lambda,\theta,\omega)$ adalah konstanta normalitas dengan $0 < \lambda < 1$. Disini tampak bahwa, untuk model penyusun $f_{SPD}(x,\lambda,\theta,\omega)$, $f_1(x,\theta)$ dan/atau $f_2(x,\omega)$, yang kompleks maka $C(\lambda,\theta,\omega)$ akan semakin sulit untuk dihitung.

Karena, $f_1(x,\theta)$ dan $f_2(x,\omega)$ merupakan fungsi dari x, sehingga $C(\lambda,\theta,\omega)$ dapat diperoleh dengan cara sebagai berikut:

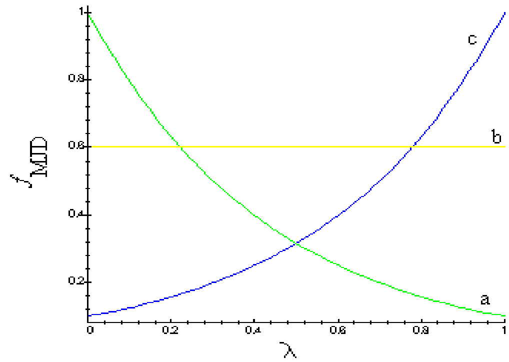
$$C\left(\lambda,\theta,\omega\right) = \left[\int_{-\infty}^{\infty} f_1^{\lambda}\left(x,\theta\right) f_2^{1-\lambda}\left(x,\omega\right) dx\right]^{-1} \quad (2)$$

Karena $C(\lambda,\theta,\omega)$ adalah suatu besaran yang konstan, maka persamaan (1) dapat dituliskan sebagai bentuk proporsional sebagai berikut :

$$f_{SPD}\left(x,\lambda,\theta,\omega\right) \propto f_1^{\lambda}(x,\theta)\, f_2^{1-\lambda}(x,\omega) \quad (3)$$

Secara umum, berdasarkan pada persamaan (1), sifat density ini dapat dituliskan sebagai berikut :

1. $f_{SPD}(x,\lambda,\theta,\omega)$ adalah sebuah distribusi probabilitas (*pdf*) dengan konstanta normalitas $C(\lambda,\theta,\omega)$ seperti dalam persamaan (2).
2. Jika λ mendekati nol, maka $f_{SPD}(x,\lambda,\theta,\omega)$ akan didominasi oleh $f_2(x,\omega)$, sedangkan jika λ mendekati satu, maka $f_{SPD}(x,\lambda,\theta,\omega)$ akan didominasi oleh $f_1(x,\theta)$. Dominasi ini menunjukkan bahwa data lebih cenderung mendekati distribusi dominatornya, yaitu $f_1(x,\theta)$ atau $f_2(x,\omega)$. Hal ini dapat dilihat pada Gambar 1.



Gambar 1 : Plot SPD dengan nilai $f_1(x,\theta)$ dan $f_2(x,\omega)$ tertentu antara nol dan satu. Garis (a) menggambarkan saat $f_1(x,\theta)=1$ dan $f_2(x,\omega)=0.1$, gambar (b) menggambarkan saat $f_1(x,\theta)=0.1$ dan $f_2(x,\omega)=1$, dan gambar (c) menggambarkan saat $f_1(x,\theta) = f_2(x,\omega)=0.6$.

3. Jika $f_1(x,\theta) = f_2(x,\omega)$, maka kekuatan dominasi antara $f_1(x,\theta)$ dan $f_2(x,\omega)$ adalah sama besar dan nilai $C(\lambda,\theta,\omega)$ adalah satu.
4. Fungsi likelihood dan log-likelihood dari $f_{SPD}(x,\lambda,\theta,\omega)$ dapat dituliskan sebagai berikut :

$$L_{f_{SPD}} = \left(C(\lambda,\theta,\omega)\right)^n \prod_{i=1}^{n} f_1^{\lambda}(x_i,\theta)\, f_2^{1-\lambda}(x_i,\omega) \tag{4}$$

dan

$$\log\left(L_{f_{SPD}}\right) = k + \sum_{i=1}^{n}\left[\lambda \log\left(f_1(x_i,\theta)\right) + (1-\lambda)\log\left(f_2(x_i,\omega)\right)\right] \tag{5}$$

Dimana nilai konstanta k adalah logaritma dari konstanta normalitas, $C(\lambda,\theta,\omega)$. Jika $f_1(x,\theta) = f_2(x,\omega)= f(x)$ sebagai kasus khusus, maka persamaan (5) akan menjadi :

$$\begin{aligned}\log\left(L_{f_{SPD}}\right) &= k + (\lambda + (1-\lambda))\sum_{i=1}^{n}\log\left(f(x_i)\right)\\ &= k + \log(L)\end{aligned} \tag{6}$$

Dimana L adalah fungsi likelihood densitas $f(x)$.

2.2 SPD dengan lebih dari dua fungsi penyusun

Jika terdapat m model dalam SPD, maka bentuk umum dari model SPD dapat dituliskan sebagai berikut :

$$f_{SPD}(x,\Lambda,\Theta)=C(\Lambda,\Theta)\prod_{i=1}^{m} f_i^{\lambda_i}(x,\theta_i)$$

$$\propto \prod_{i=1}^{m} f_i^{\lambda_i}(x,\theta_i) \qquad (7)$$

Dengan $\Lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m), \Theta = (\theta_1, \theta_2, \ldots, \theta_m)$, $0 \le \lambda_i \le 1$, $\Sigma_{i=1..m}\lambda_i = 1$, dan $C(\Lambda,\Theta)$ dihitung sesuai dengan persamaan (2). Model likelihoodnya adalah

$$L_{f_{SPD}} = (C(\Lambda,\Theta))^n \prod_{j=1}^{n}\prod_{i=1}^{m} f_i^{\lambda_i}(x_j,\theta_i) \qquad (8)$$

Jika semua m model tersebut sama, maka log-likelihoodnya akan dapat dituliskan seperti persamaan (6) setelah pemisahan suku $\Sigma_{i=1..m}\lambda_i = 1$.

2.3 Implementasi Pemilihan model dengan SPD

Pemilihan model yang tepat untuk suatu data dengan menggunakan SPD dilakukan dengan pendekatan model Bayesian. Penaksiran parameter setiap model dalam SPD yang dilanjutkan dengan penentuan nilai dominator λ, dilakukan melalui model marginal setiap parameternya. Model marginal yang digunakan disini adalah model marginal penuh, yaitu suatu model marginal suatu parameter dalam SPD dengan semua parameter yang lain dalam SPD sudah diketahui nilainya. Dengan menentukan distribusi prior yang sesuai untuk setiap parameter model ((Iriawan,1999) dan (Carlin dan Chib, 1995)), maka taksiran distribusi posterior setiap parameter akan dapat diperoleh dari marginalnya.

Anggap jika dalam SPD terdapat dua model penyusun, seperti dalam persamaan (1), dengan masing-masing terdapat sebanyak p dan q parameter, maka akan didapatkan sebanyak (p+q+1) model marginal penuh. Dengan menggunakan MCMC, khususnya Gibbs sampler, dalam pemilihan model ini akan dibutuhkan sebanyak (p+q +1) tahap penaksiran. Dimana satu tahap tambahan tersebut adalah penaksiran nilai dominator λ. Akhirnya, dalam N kali iterasi dalam pembangkitan dominator λ, rasio antara banyaknya model yang terpilih, M_g, terhadap seluruh model yang dibangkitkan, N, akan dapat diperoleh taksiran Bayes faktor yang digunakan untuk memilih model mana yang paling relevan untuk merepresentasikan datanya. Taksiran Bayes faktor ini dapat diperoleh dengan menggunakan pendekatan nilai probabilitas berikut :

$$\hat{p}(M_g \mid Y) = \frac{\text{Banyaknya model sesuai}\,\lambda\ \ \text{yang diperoleh}}{\text{Banyaknya iterasi pembangkitan model}} \qquad (9)$$

3. Contoh Numerik

3.1 Pemilihan Distribusi Data

Contoh penerapan pertama yang akan dibahas dalam makalah ini adalah mengenai pemilihan distribusi suatu data. Data dibangkitkan dari distribusi Exponensial dan kemudian diuji dengan menggunakan SPD untuk dibandingkan dengan Log-Normal

(O'Hagan, 1995). Algoritma MCMC untuk mengimplementasikan SPD ini dibutuhkan sebanyak empat tahapan, yaitu : satu tahap membangkitan taksiran parameter eksponensial, dua tahap membangkitkan parameter Log-Normal, dan satu tahap membangkitkan taksiran dominator SPD, λ. Dengan menggunakan kriteria pemilihan bahwa model pertama dalam SPD pada iterasi ke-*g*, $M_1^{(g)}$, jika $\lambda^{(g)} < 0{,}5$ dan dengan melakukan iterasi sebanyak 10000 kali dalam MCMC dimana 5000 iterasi dianggap 'burn-in', maka diperoleh hasil seperti dalam Tabel 1. Dari hasil tersebut dapat dilihat bahwa MCMC pada SPD di atas terbukti mampu menunjukkan bahwa suatu data berasal dari suatu distribusi yang benar.

Tabel 1: Banyaknya model yang dibangkitan, Bayes faktor dan 95% Interval Kepercayaan dari Bayes faktor

Distribusi	Bangkitan Model		Bayes	95% CI dari
Data	Exp	LN	faktor	Bayes faktor
Exp(0.2)	4908	92	53.3478	(42,3208;62,6554)
Exp(1)	4345	655	6.6336	(3,71458;7,8068)
Exp(10)	4693	307	15.2866	(12,5150;18,2158)
Exp(20)	4980	20	249	(216,3307;321,9787)
LN(0,0.2)	1192	3808	0.313	(0,2462;0,3434)
LN(2,3)	24	4976	0.0048	(0,0;0,0216)
LN(4,4)	60	4940	0.0121	(0,0;0,0414)
LN(5,4)	21	4979	0.0004	(0,0;0,0194)
exp(N(0,1))	281	4719	0.0597	(0,0201;0,0992)

3.2 Model Regresi tidak bersarang

Carlin dan Chib (1995) melakukan pemiliham model terbaik dari dua model yang berbeda pada 42 data spesimen suatu 'radiata pine compressive strength' (Williams, 1959). Data ini ditampilkan pada Tabel 2. Kedua model yang dibandingkan adalah :

$$f_1 = M_1 : y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1..42 \tag{10}$$

$$f_2 = M_2 : y_i = \gamma + \theta z_i + \xi_i, \qquad i = 1..42 \tag{11}$$

dimana ε_i dan ξ_i bersifat IIDN$(0,\sigma^2)$ dan IIDN$(0,\tau^2)$,

y_i adalah kekuatan tekanan maksimum,

x_i adalah densitas specimen,

z_i adalah densitas x_i yang telah direaksikan dengan zat tertentu.

Fungsi likelihood dari SPD untuk kedua model adalah sebagai berikut :

$$\begin{aligned} L_{f_{SPD}} &= \left(\frac{C(\lambda,\omega_1,\omega_2)}{2\pi\sigma^{\lambda}\tau^{(1-\lambda)}} \right)^n \\ &= \exp\left(-\sum_{i=1}^{n} \left[\frac{\lambda}{2}\left(\frac{y_i-\alpha-\beta x_i}{\sigma} \right)^2 + \left(\frac{1-\lambda}{2} \right)\left(\frac{y_i-\gamma-\theta z_i}{\tau} \right)^2 \right] \right) \end{aligned} \tag{12}$$

dimana $C(\lambda,\omega_1,\omega_2)$ dihitung seperti dalam persamaan (2) dan
$\omega_1=(\alpha,\beta,\sigma)$ dan $\omega_2=(\gamma,\theta,\tau)$.

Tabel 2 : Data 'Radiata pine compressive strength'

No	*y*	*x*	*z*	*No*	*y*	*x*	*Z*
1	3040	29.2	25.4	22	3840	30.7	30.7
2	2470	24.7	22.2	23	3800	32.7	32.6
3	3610	32.3	32.2	24	4600	32.6	32.5
4	3480	31.3	31.0	25	1900	22.1	20.8
5	3810	31.5	30.9	26	2530	25.3	23.1
6	2330	24.5	23.9	27	2920	30.8	29.8
7	1800	19.9	19.2	28	4990	38.9	38.1
8	3110	27.3	27.2	29	1670	22.1	21.3
9	3160	27.1	26.3	30	3310	29.2	28.5
10	2310	24.0	23.9	31	3450	30.1	29.2
11	4360	33.8	33.2	32	3600	31.4	31.4
12		21.5	21.0	33	2850	26.7	25.9
13	3670	32.2	29.0	34	1590	22.1	21.4
14	1740	22.5	22.0	35	3770	30.3	29.8
15	2250	27.5	23.8	36	3850	32.0	30.6
16	2650	25.6	25.3	37	2480	23.2	22.6
17	4970	34.5	34.2	38	3570	30.3	30.3
18	2620	26.2	25.7	39	2620	29.9	23.8
19	2900	26.7	26.4	40	1890	20.8	18.4
20	1670	21.1	20.0	41	3030	33.2	29.4
21	2540	24.1	23.9	42	3030	28.2	28.2

Sehingga MCMC pada SPD di atas akan mempunyai tujuh level, yaitu terdiri dari tiga level untuk proses penaksiran pameter model Normal dari kedua penyusunnya; yaitu $\omega_1=(\alpha,\beta,\sigma)$ dan $\omega_2=(\gamma,\theta,\tau)$ dan satu level untuk penaksiran parameter indikator SPD, λ. Untuk mendapatkan posterior tiap parameter SPD tersebut digunakan distribusi prior untuk (α,β) dan (γ,θ) adalah Normal dengan mean dan varians : $((3000; 185)^T, (10^6;10^4))$ dan distribusi prior untuk σ dan τ digunakan distribusi Inverse-Gamma dengan mean dan varians 300^2. Parameter distribusi prior ini didapatkan dari data asli x_i dan z_i dalam Tabel 2 setelah distandarkan pada pusat rata-ratanya (Carlin dan Chib, 1995). Sedangkan, distribusi prior untuk λ digunakan distribusi Beta(a,a) atau Beta yang simetri pada range (0, 1).

Selanjutnya, dengan menggunakan karakteristik nilai λ sesuai dengan MCMC dalam contoh pertama, sub-bagian 3.1, yaitu $\lambda^{(g)} < 0{,}5$ menunjukkan model pertama mendominasi SPD dan MCMC dilakukan 10000 iterasi dimana 5000 iterasi awalnya digunakan sebagai 'burn in', maka diperoleh pembangkitan model sebesar 221 untuk model pertama dan 4779 untuk model kedua. Banyaknya bangkitan model ini senilai dengan taksiran Bayes faktor sebesar 21.6244. Selanjutnya untuk memperoleh gambaran kisaran nilai Bayes faktor ini, 1000 run MCMC diatas dijalankan sebanyak 100 kali dan diperoleh taksiran rata-rata Bayes faktor 21.6176 yang berada dalam 95% interval kepercayaan antara 17.4555 dan 25.7796. Nilai taksiran Bayes faktor ini menunjukkan suatu bukti yang kuat bahwa data dalam Tabel 2 tersebut lebih cenderung akan lebih baik jika dimodelkan ke dalam model persamaan (11) dari pada model dalam persamaan (10) (Kass dan Raftery, 1995).

4. Kesimpulan

Makalah ini telah mambahas SPD dan penghitungan taksiran nilai Bayes faktor serta penggunaannya dalam pemilihan model. Bukti berlakunya metode ini telah diujikan pada dua contoh pemilihan distribusi data dan pada pemilihan model regresi tidak bersarang yang masing-masing mempunyai error dengan asumsi Normal. Seperti halnya dalam penggunaan likelihood ratio dalam pemilihan model, dengan mengacu pada Kass dan Raftery (1995), taksiran nilai Bayes faktor dapat digunakan sebagai alat bantu pemilihan model dengan SPD yang diimplementasikan menggunakan MCMC. Pembahasan lebih mendalam mengenai pemilihan model tidak bersarang dengan SPD ini dapat dilihat pada Iriawan (1999) dengan asumsi bentuk errornya adalah sembarang atau neo-Normal, MSNBurr(k,α,μ,ϕ).

Daftar Pustaka

Box, G. E. P. dan Tiao, G. C.: 1973, *Bayesian Inference in Statistical Analysis*, Reading, MA : Addison-Wesley.

Carlin, B. P. dan Chib, S.: 1995, Bayesian model choice via Markov Chain Monte Carlo methods, *Journal of the Royal Statistical Society, Ser. B* **57**(3), 473-484.

Carlin, B. P., Polson, N. G. dan Stopper, D. S.: 1992, A monte carlo approach to nonnormal and nonlinear state-space modelling, *Journal of the American Statistical Association* **87**(418), 493-500.

Carota, C., Parmigiani, G. dan Polson, N. G.: 1996, Diagnostic measures for model criticism, *Journal of the American Statistical Association* **91**(434), 753-762.

Casella, G. dan George, E. I.: 1992, Explaining Gibbs sampler, *Journal of the American Statistical Association* **46**(3), 167-174.

Chib, S.: 1995, Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association* **90**(432), 1313-1321.

Cox, D. R.: 1962, Further results on tests of separate families of hypotheses, *Journal of the Royal Statistical Society Ser. B* **24**(2), 406-424.

Gelfand, A. E. dan Dey, D. K.: 1994, Bayesian model choice : Asymptotics and exact calculations, *Journal of the Royal Statistics Society Ser. B.* **56**(3), 501-514

Gelfand, A. E., Hills, S. E., Racine-Poon, A. dan Smith, A. F. M.: 1990, Illustration of Bayesian inference in normal data models using Gibbs sampling, *Journal of the American Statistical Association* **85**(412), 972-985.

Geman, S. dan Geman, D.: 1984, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6), 721-741.

Gilks, W. R., Clayton, D. G., Spiegelhalter, D. J., Best, N. G., McNeil, A. J., Sharples, L. D. dan Kirby, A. J.: 1993, Modeling complexity : Applications of Gibbs sampling in medicine, *Journal of the Royal Statistics Society*, *Ser. B* **55**(1), 39-52.

Iriawan, N : 1999, *Computational Intensive Approaches to Inference in Neo-Normal Linear Models*, Ph.D. Thesis.

Kass, R. E. dan Raftery, A. E.: 1995, Bayes factors, *Journal of the American Statistical Association* **90**(430), 773-795.

Newton, M. A. dan Raftery, A. E.: 1994, Approximate Bayesian inference with the weighted likelihood bootstrap, *Journal of the Royal Statistical Society*, *Ser. B* **56**(1), 3-48.

O'Hagan, A.: 1995, Fractional Bayes factors for model selection, *Journal of the Royal Statistical Society, Ser. B.* **57**(1), 99-138.

Wakefield, J. C., Smith, A. F. M., Racine-Poon, A. dan Gelfand, A. E.: 1994, Bayesian analysis of linear and non-linear population models by using the Gibbs sampler, *Applied Statistics* **43**(1), 201-221.

Williams, E. J.: 1959, *Regression Analysis*, John Wiley & Sons, New York.

Zeger, S. L. dan Karim, M. R.: 1991, Generalized linear models with random effects; a Gibbs sampling approach, *Journal of the American Statistical Association* **86**(413), 79-86.