Advances in Applied Mathematical Biosciences.
ISSN 2248-9983 Volume 4, Number 2 (2013), pp. 63-68

http://www.irphouse.com

Diagnose of Primary Tumor Cancer using Markov Chain Monte-Carlo Convergence Model

Dr. K. Vaikundamoorthy

Associate Professor & Head,
Department of Mathematics, Kalasalingam University,
Krishnankovil, Tamilnadu, India,
E-mail id: kvmorthy@yahoo.com.

Abstract

Maximum probability of existence of cancer in human bodies is normally diagnosed very late, so that, it is highly cumbersome for physicians to cure. Reliability in predicting cancer at initial stage is always needed, so that curing and medical recovery is possible. In this paper, an investigation was made to diagnose the presence of primary tumor using MCMC Convergence model. The MCMC procedure is used here to carry out the analysis which is most efficient on a wide range of complex Bayesian statistical models. The analysis was carried out using version 18 of SPSS AMOS software. Totally, 18 components were considered for the diagnosis from the primary tumor samples of 725 patients. Patients having primary tumor were analysed considering various factors such as class, age, sex, degree of life, etc. using mathematical modeling techniques. The maximum likelihood estimators (MLEs) of the parameters were derived and assessed their performance through a Monte Carlo simulation study. From the collected information the values of convergence for likelihood of each components of primary tumor has been identified and the results presented.

Key words*:* Monte Carlo Simulation, Modelling, Cancer, Convergence

1. INTRODUCTION

Recent developments in science and technology in the past few centuries has made it necessary to apply mathematical methods to real-life problems arising from different fields – be it Science, Finance, Management etc. With the advent of computational power of digital computers and computing methods, the use of Mathematics in solving real-world problems has become widespread, especially for handling of lengthy and complicated problems.

The process of translation of a real-life problem into a mathematical form can give a better representation and solution of certain critical problems. Markov chain Monte Carlo (MCMC) methods are a class of algorithms used for sampling from probability distributions based on constructing a Markov chain as its equilibrium distribution. After a large number of steps / iteration, it is used as a sample of the desired distribution. Based on the number of steps, the quality of the sample improves. The number of steps required is determined based on the convergence to the stationary distribution within an acceptable error.

2. LITERATURE REVIEW

Studies on statistical modelling have been reported by various researchers such as models through Air pollution data, [Cowles et al (2002)] A single MCMC chain, [Sylwestrowicz (1982)], [Adams et al (1996)], [Rossini et al (2003)], [Rosenthal (2000)], [Wilkinson (2005)] Spatial statistical modeling [Whiley and Wilson (2004)], [Blackford et al, (1997)], [Neal (2003)] and Bayesian Spatiotemporal Geo-statistical Model. Research works of Bayesian Analysis of Stochastic Models were carried out in Single Molecule Biophysics. Recent technological advances have allowed scientists to follow a biochemical process on a single molecule basis, unlike traditional macroscopic experiments. These raised many challenging data-analysis problems and called for a sophisticated statistical modeling and inference effort.

In this paper, an investigation was made to diagnose the patients of primary tumor cancer using MCMC Convergence model. 18 components were considered for the diagnosis from the primary tumor samples of 725 patients. Patients having primary tumor cancer were analysed considering various components such as class, age, sex, degree of life, etc. using mathematical modeling techniques. The maximum likelihood estimators (MLEs) of the parameters were derived and assessed their performance through a Monte Carlo simulation study. The Bayesian statistical model of MCMC procedure is used in this analysis, which enables us to carry out analysis on a wide range of complex data. The analysis was carried out using version 18 of SPSS AMOS software.

3. DATA PROCESSING AND ANALYSIS

Amos provides several diagnostics that help anyone check convergence. The patients having primary tumor components were taken into account. Class: lung, head and neck, esophagus, thyroid, stomach, duodenum and 5m. int, colon, rectum, anus, salivary glands, pancreas, gallbladder, liver, kidney, bladder, testis, prostate, ovary, corpus, uteri, cervix uteri, vagina and breast. Age: <20, 20-59, greater than or equal to 60. Sex: Male, Female. Histologic-type (epiddermoid, adeno, anaplastic). Degree of life (well, fair, poor), bone (yes, no), bone-marrow (yes, no), lung (yes, no), pleura (yes, no), peritoneum (yes, no). liver (yes, no). brain (yes, no). skin (yes, no). neck (yes, no). supraclavicular (yes, no), axillar (yes, no), mediastinum (yes, no), abdominal (yes, no). All types of the patients in all types of primary tumor were strongly diagnosed through the model.

Bayesian analysis requires estimation of explicit means and intercepts. Before performing any Bayesian analysis in Amos, we have to first tell Amos to estimate means and intercepts. Amos displays Estimates, Scalar Estimates, Maximum Likelihood Estimates, and Regression Weights. F1-F2 diagram is then obtained after analyzing the tables in Amos and given in Fig. 1. In F1-F2 diagram, F2 contains class, age, sex and type while F1 includes all other components for analyzing primary tumor cancer. Regression weights, Intercepts, Co-variances and Variance are predicted by AMOS software and the results are presented in Table 1.

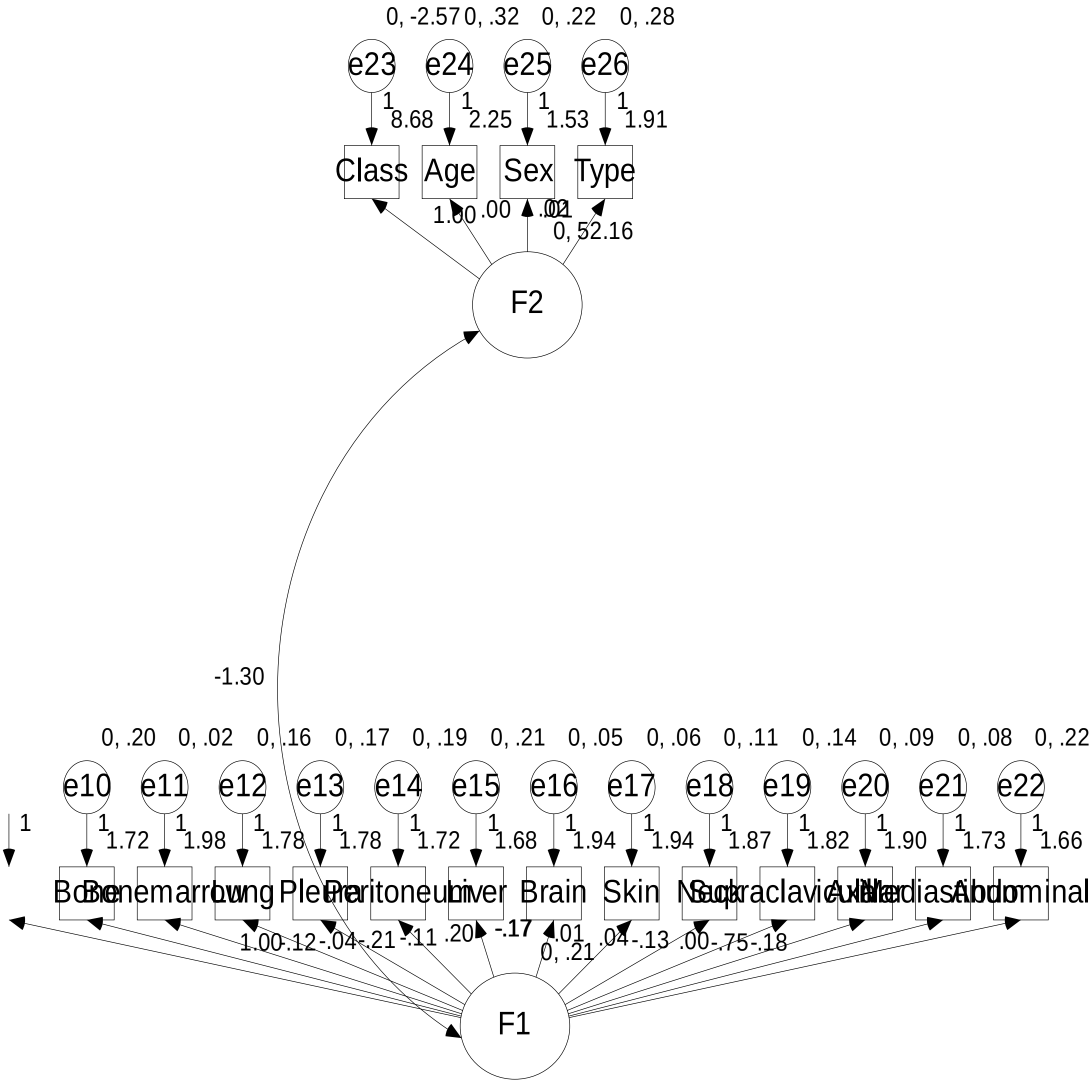

Fig. 1: F1-F2 Diagram

Table 1: Results on Convergence Analysis (C.S.)

	Mean	S.E.	S.D.	C.S.	Median	95% Lower bound	95% Upper bound	Skewness	Kurtosis	Min	Max
Regression weights											
Bone<--F1	-0.15	0.003	0.092	1.001	-0.147	-0.338	0.026	-0.172	0.297	-0.577	0.231
Bonemarrow<--F1	-0.053	0.001	0.03	1	-0.051	-0.117	0.002	-0.3	0.553	-0.196	0.065
Lung<--F1	-0.305	0.009	0.109	1.003	-0.295	-0.548	-0.118	-0.521	0.503	-0.799	0.055
Pleura<--F1	-0.166	0.007	0.1	1.002	-0.16	-0.379	0.015	-0.287	0.194	-0.607	0.2
Peritoneum<--F1	0.244	0.004	0.102	1.001	0.238	0.056	0.467	0.391	0.737	-0.142	0.718
Liver<--F1	-0.217	0.005	0.108	1.001	-0.21	-0.451	-0.026	-0.42	0.468	-0.702	0.194
Brain<--F1	-0.219	0.004	0.062	1.002	-0.213	-0.359	-0.115	-0.568	0.432	-0.531	-0.038
Skin<--F1	0.02	0.002	0.05	1.001	0.018	-0.073	0.124	0.326	0.778	-0.158	0.328
Neck<--F1	0.042	0.002	0.073	1	0.041	-0.103	0.186	0.025	0.285	-0.298	0.4
Supraclavicular <--F1	-0.187	0.007	0.089	1.003	-0.179	-0.387	-0.032	-0.503	0.629	-0.592	0.174
Axillar<--F1	-0.023	0.004	0.063	1.002	-0.021	-0.155	0.098	-0.128	0.321	-0.312	0.301
Mediastinum<--F1	-1.152	0.042	0.346	1.007	-1.096	-1.946	-0.626	-0.686	0.081	-2.36	-0.362
Abdominal<--F1	-0.273	0.011	0.137	1.003	-0.262	-0.576	-0.032	-0.54	0.914	-0.879	0.311
Age<--F2	0.004	0.001	0.009	1.005	0.002	-0.01	0.029	1.08	1.843	-0.024	0.052
Sex<--F2	0.045	0.003	0.021	1.012	0.039	0.023	0.102	1.514	1.742	0.013	0.119
Type<--F2	0.012	0.001	0.01	1.008	0.01	-0.002	0.037	1.309	2.557	-0.021	0.061
Intercepts											
Difference	2.015	0.001	0.045	1	2.015	1.927	2.102	0.007	-0.067	1.822	2.215
Bone	1.723	0	0.025	1	1.723	1.674	1.77	-0.033	0.033	1.616	1.837
Bonemarrow	1.979	0	0.008	1	1.979	1.964	1.994	0.027	-0.002	1.949	2.01
Lung	1.778	0.001	0.023	1	1.778	1.733	1.822	-0.044	0.009	1.682	1.871
Pleura	1.778	0	0.023	1	1.778	1.732	1.822	-0.058	0.012	1.676	1.868
Peritoneum	1.72	0.001	0.025	1	1.72	1.671	1.768	-0.068	-0.085	1.621	1.807
Liver	1.679	0	0.026	1	1.679	1.628	1.729	-0.004	0.105	1.574	1.784
Brain	1.938	0	0.013	1	1.938	1.912	1.964	0.015	-0.065	1.886	1.992
Skin	1.941	0	0.013	1	1.941	1.915	1.966	-0.036	-0.014	1.889	1.995
Neck	1.871	0	0.018	1	1.871	1.834	1.907	-0.039	0.032	1.791	1.954
Supraclavicular	1.821	0	0.021	1	1.821	1.78	1.862	-0.009	-0.027	1.719	1.899
Axillar	1.902	0	0.016	1	1.902	1.869	1.934	-0.032	0.023	1.84	1.966
Mediastinum	1.728	0.001	0.024	1	1.729	1.68	1.776	-0.035	0.029	1.613	1.825
Abdominal	1.661	0.001	0.026	1	1.661	1.611	1.713	0.062	0.046	1.551	1.782
Class	8.67	0.007	0.389	1	8.675	7.902	9.424	-0.023	-0.005	7.185	10.159
Age	2.248	0.001	0.032	1	2.248	2.186	2.31	-0.015	-0.036	2.12	2.375
Sex	1.524	0.001	0.028	1	1.525	1.47	1.578	-0.018	0.009	1.406	1.639
Type	1.912	0	0.029	1	1.912	1.856	1.969	-0.004	0.064	1.78	2.035
Covariances											
F2<->F1	-0.797	0.039	0.362	1.006	-0.765	-1.585	-0.176	-0.522	0.304	-2.576	0.035
Variances											
F1	0.121	0.005	0.051	1.004	0.114	0.046	0.236	0.773	0.623	0.028	0.405
F2	32.592	1.275	10.811	1.007	33.222	11.75	51.453	-0.165	-0.697	5.926	64.589
e9	0.568	0.003	0.058	1.002	0.568	0.454	0.682	0.002	0.002	0.332	0.796
e10	0.201	0	0.016	1	0.2	0.173	0.235	0.305	0.229	0.146	0.282
e11	0.02	0	0.002	1	0.02	0.017	0.024	0.289	0.161	0.015	0.028
e12	0.165	0	0.013	1	0.165	0.141	0.193	0.271	0.037	0.116	0.226
e13	0.172	0	0.014	1	0.171	0.147	0.201	0.269	0.07	0.121	0.235
e14	0.198	0	0.016	1	0.198	0.169	0.231	0.257	0.12	0.134	0.271
e15	0.217	0	0.017	1	0.216	0.186	0.253	0.407	0.379	0.161	0.314
e16	0.054	0	0.004	1	0.054	0.046	0.063	0.309	0.236	0.038	0.075
e17	0.056	0	0.004	1	0.056	0.048	0.066	0.271	0.08	0.042	0.077
e18	0.115	0	0.009	1	0.114	0.099	0.132	0.215	0.023	0.082	0.16
e19	0.147	0	0.012	1	0.146	0.126	0.171	0.34	0.177	0.106	0.202
e20	0.089	0	0.007	1	0.089	0.077	0.104	0.34	0.192	0.066	0.124
e21	0.064	0.003	0.03	1.004	0.065	0.007	0.121	0.03	-0.423	0	0.186
e22	0.22	0	0.018	1	0.219	0.188	0.258	0.38	0.319	0.152	0.31
e23	17.759	1.237	10.304	1.007	16.942	1.234	39.365	0.376	-0.549	0	49.859
e24	0.327	0	0.026	1	0.325	0.281	0.38	0.353	0.369	0.239	0.48
e25	0.197	0.004	0.03	1.008	0.201	0.121	0.244	-0.905	1.036	0.073	0.289
e26	0.283	0.001	0.022	1	0.281	0.242	0.33	0.307	0.09	0.209	0.392

4. RESULTS AND DISCUSSIONS

On the toolbar of the Bayesian SEM window, AMOS presented a convergence value (C.S) of 1.0083. This is an overall convergence based on the statistical analysis. Each time the screen refreshes, Amos updates the *C.S.* for each parameter in the summary table; the *C.S.* value on the toolbar is the largest of the individual *C.S.* values. The *C.S.* compares the variability within parts of the analysis sample to the variability across these parts. By this standard, the maximum *C.S.* of 1.0083 is not small enough, then the fitness displays an unhappy face.

5. CONCLUSIONS

From all the collected information and the statistical analysis carried out using AMOS, each component of primary tumor has the maximum C.S. which is strictly less than 1.0083. Thus, the patients having primary tumor cancer are strongly diagnosed through the model.

6. REFERENCES

[1] Adams, N. M., Kirby, S. P. J., Harris, P., and Clegg, D. B. (1996), A review of parallel processing for statistical computation, *Statistics and Computing*, 6, pp. 37–49.

[2] Blackford, L. S., Choi, J., Cleary, A., D'Azevedo, E., Demmel, J., Dhillon, I., Dongarra, J., Hammarling, S., Henry, G., Petitet, A., Stanley, K., Walker, D., and Whaley, R. C. (1997), *ScaLAPACK Users' Guide*, Philadelphia, PA: Society for Industrial and Applied Mathematics.

[3] Cowles, M. K., Zimmerman, D. L., Christ, A., and McGinnis, D. L. (2002), Combining snow water equivalent data from multiple sources to estimate spatio-temporal trends and compare measurement systems, *Journal of Agricultural, Biological, and Environmental Statistics* 7, 536–557.

[4] Neal, R. M. (2003), Slice sampling, *The Annals of Statistics*, 31, 705–767.

[5] Rosenthal, J. S. (2000), Parallel computing and Monte Carlo algorithms Far East, *Journal of Theoretical Statistics*, 4, 207–236.

[6] Rossini, A., Tierney, L., and Li, N. (2003), *Simple parallel statistical computing in R*, Working paper, UW Biostatistics.

[7] Sylwestrowicz, J. D. (1982), Parallel processing in statistics. In H. Caussinus, P. Ettinger, and R. Tomassone (eds.), COMPSTAT 1982, *Proceedings in Computational Statistics*, pp. 131–136, Physica-Verlag Ges.m.b.H. van de Geijn, R. A. 1997. Using PLAPACK. The MIT Press.

[8] Wilkinson, D. J. (2005), *Parallel Bayesian computation*. In E. J. Kontoghiorghes (ed.), Hand- book of Parallel Computing and Statistics, pp. 481–512, Marcel Dekker/CRC Press.

[9] Whiley, M. and Wilson, S. P. (2004), Parallel algorithms for Markov chain Monte Carlo methods in latent spatial Gaussian models, *Statistics and Computing*, 14, 171–179.

[10] Liu, C. (2001), Discussion on The Art of Data Augmentation, *Journal of Computational and Graphical Statistics*.
[11] Box, G.E.P. and Jenkins G.M. (1976), *Time series: Analysis, Forecasting and Control* (Holden Day, San Francisco).
[12] Monahan, J. (1984), Full Bayesian analysis of ARMA time series models, *Journal of Econo- metrics*, 21, 307-331.