Fakulti Teknologi Maklumat dan Sains Kuantitatif ISSN:1823-0822 Jilid 7, Bil. 1, 2005

UNIVERSITI TEKNOLOGI MARA

JURNAL TEKNOLOGI MAKLUMAT DAN SAINS KUANTITATIF

Kandungan **Muka Surat**

Different Types of Interpolations for Solving Delay Differential Equations using Explicit Runge-Kutta Method 1
Fudziah Ismail, Ang San Lwin, Mohamed Suleiman

A Preliminary Study on the Collaborative Use of Statistical Modeling in a GIS Study Of Asthmatic Morbidity 9
Mohammad Said Zainol, Sayed Jamaluddin S.Ali, Zainal Mat Saat

Implementing Slicing Technique on JPEG-File-Its Impact on the Download Time 19
Fakhrul Hazman Yusoff, Anita Mohd Yasin, Rozianawaty Osman

Quantifying Consensus on Women's Roles using Fuzzy Logic 29
Puzziawati Ab Ghani, Abdul Aziz Jemain

Early Identification of Low Employability Graduate in Malaysia: The use of Proportional Hazard Model 41
Lim Hock-Eam

E-Service Quality: Malaysian Perceptions 49
Noor Habibah Arshad, Norjansalika Janom, Isnainy Mohd Idris

Research Performance Evaluation using Data Envelopment Analysis (DEA) 59
Norshahida Shaadan

Universiti Teknologi MARA

LEMBAGA PENYUNTING

Penasihat
Prof. Madya Dr.Adnan Ahmad
Dekan Fakulti Teknologi Maklumat dan Sains Kuantitatif

Ketua Penyunting
Prof. Dr. Mohd Sahar Sawiran

Penyunting
Prof. Madya Dr. Daud Mohamad
Prof. Madya Dr. Mazani Manaf
Prof. Madya Dr. Saadiah Yahaya
Prof. Madya Dr. Yap Bee Wah
Prof. Madya Ooi Hee Tang
Prof. Madya Dalialah Abd. Ghani
Mohd Hanafi Tumin ASA

Pengurusan Penerbitan
Prof. Madya Dr. Mazani Manaf
Puan Zahrah Hj Abdul Razak

Penyunting Luaran
Prof. Madya Dr. Khairuddin Omar
Jabatan Sains dan Pengurusan Sistem
Fakulti Teknologi dan Sains Maklumat, UKM Bangi
E-mail: kjts@ftms.ukm.my

Prof. Madya Dr. Tahir Ahmad
Jabatan Matematik
Fakulti Sains, UTM Skudai
E-mail: tahir_@hotmail.com

Prof. Madya Dr Hapsah Midi
Jabatan Matematik
Faculty of Science and Enviromental Studies,
UPM Serdang
E-mail: habshah@fsas.upm.edu.my

Dasar Penerbitan:
Jurnal Teknologi Maklumat dan Sains Kuantitatif diterbitkan oleh Fakulti Teknologi Maklumat dan Sains Kuantitatif, Universiti Teknologi MARA. Sumbangan penulisan adalah berkaitan dengan teori, amali, metodologi serta falsafah aspek-aspek kefenomenan dan epistemologi dalam sains matematik dan pengkomputeran. Tujuan utama jurnal ini adalah untuk mengenengahkan bahan/karya yang menunjuk dan mempersembahkan keselarasan serta keharmonian dalam teknologi maklumat dan sains kuantitatif. kandungan makalah yang dimuatkan dalam jurnal ini tidak semestinya menderminkan pandangan dan pendirian rasmi jurnal ini.

Alamatkan semua sumbangan kepada:
Urusetia Panel Penyunting,
Jurnal Teknologi Maklumat dan Sains Kuantitatif
Fakulti Teknologi Maklumat dan Sains Kuantitatif
UiTM Shah Alam

Tel **03-5543 5329**
Faks **03-5543 5501**
e-mail mshahar@tmsk.uitm.edu.my

DARI MEJA KETUA PENYUNTING

Alhamdulillah, dapat kita terbitkan Jurnal Teknologi Maklumat dan Sains Kuantitatif Jilid 7, Bil.1, 2005. Saya rasa pencinta ilmu menanti-nanti terbitan kali ini.

Seperti biasa jurnal terbitan sesuatu tahun itu, hanya dapat dihantar untuk percetakan dua atau tiga bulan berikutnya. Kadangkala, penulis yang telah menghantar balik artikel yang telah diwasitkan itu tertunggu-tunggu juga adakah artikelnya diterbitkan kali ini. Sememangnya pihak penyunting mengamalkan prinsip giliran FIFO (first in first out), tetapi kadangkala ianya tidak boleh dilakukan. Ini kerana sesuatu bidang pengkhususan itu mempunyai dua atau tiga artikel sekaligus. Jadi pihak penyunting berkemungkinan akan melewatkan salah satu daripada artikel sebidang itu kemudian. Justeru itu, giliran FIFO masih dilakukan dalam bidang yang sama.

Dalam keluaran yang lepas, saya ada mengatakan bahawa minat penulis akan terhakis apabila maklumbalas tentang penerimaan sesuatu artikel untuk diterbitkan itu lambat. Saya hanya boleh memberi nasihat kepada penulis supaya bersabar, sebab ini begantung kepada pewasit yang menilai itu sibuk atau tidak, sanggup atau tidak dan sebagainya. Percayalah, kesabaran itu akan menjadi kita penulis yang berdisiplin.

Akhir kata, saya harap semua penulis-penulis semasa dan yang akan datang tetap gigih untuk menulis supaya karya kita dapat dimanfaatkan oleh para ilmuwan yang lain dalam bidang kita iaitu Teknologi Maklumat dan Sains Kuantitatif

Terima kasih.

Ketua Penyunting.
Prof. Dr. Mohd Sahar Sawiran

Jurnal Tek. Maklumat & Sains Kuantitatif ISSN 1823-0822 Jilid 7,Bil 1, 2005

Early Identification Of Low Employability Graduate In Malaysia: The Use Of Proportional Hazard Model

Lim Hock-Eam
Department of Economics, Monash University, Clayton, Australia
&
Faculty of Economics, Universiti Utara Malaysia, 06010 Sintok, Kedah.
Email: lheam@uum.edu.my

Abstract

This paper illustrates the use of statistical prediction model for early identification of low employability Malaysian graduates using the proportional hazard model. The relative predicted hazard rate or probability of exit from unemployment is used to proxy the graduate's employability. The out-of-sample evaluation shows that the statistical prediction model predicts correctly 83%, 75% and 80% of the graduates that identified at bottom 15%, 20% and 25% of employability respectively. The estimated probability of exit from unemployment also suggests that nine months is the reasonable expected unemployment duration for Malaysian graduates.

Keywords: Early identification; Graduate employability; Graduate unemployment; Proportional hazard model; Statistical prediction model.

1. Introduction

The problem of graduate unemployment is gaining concern in Malaysia since the currency crisis of 1997. Despite the increasing growth of economics and million of dollars invested on the graduate re-training program, the graduate unemployment problem is persistent. In literature, there are extensive studies on the graduate unemployment problem. The determinants have been identified, such as the low English language proficiency, mismatch of skills, the increasing supply of graduates and the income while unemployed (Lim & Normizan, 2002; Roed & Zhang, 2003; Morshidi Sirat et al., 2004). Knowing the associated determinants is the first step to combat the graduate unemployment problem. The prediction of the going-to-be graduate's employability is crucial as well. This prediction enables the graduates to take appropriate remedies to improve their

employability. It also helps the authorities to allocate the limited places of their graduate re-training programs. This is consistent with the concept of quality control – rectify the problem before it occurs.

The statistical prediction model has been used to predict the risk of Sudden Infant Death Syndrome (Carpenter, 1983); the risk of reconviction of prisoner that to be considered released on parole (Copas & Marshall, 1998). Since 1990s, it applies to the early identification of jobseekers at risk of becoming long-term unemployed or exhausting the unemployment benefit. Using the UK data, Payne & Payne (2000) suggest a fixed-time model (using binary logistic regression model) to predict those with high probability of being unemployed for 12 months or more. Black et al. (2003) evaluate the statistical prediction model of unemployment benefit claimants in USA. They found that the simple linear probability model outperforms the advance nonlinear discrete choice models - the continuous dependent variable is more efficient than categorical. In Australia, the statistical prediction model has been used together with other tools to predict the level of job seeker disadvantage and targeting the appropriate assistance. Other countries are reported as beginning to experiment with the statistical prediction model for the early identification of long term unemployed (OECD, 1998).

This statistical prediction model can be used for early identification of low employability graduate in Malaysia. Besides, the predicted hazard rate also can provide hints about the reasonable expected unemployment duration of graduates. Nevertheless, to my knowledge, it is surprise that there is no attempt to use the statistical prediction model for early identification of low employability graduates in Malaysia. To fill the gap, the objective of this paper is to illustrate the use of statistical prediction model for this early identification.

This paper consists of four sections. First, the brief introduction and literature review. Second section describes the data and methodology. The third section presents the result and the final section concludes the result of this paper.

2. Data & Methodology

2.1 Data

The data used in this paper are obtained from Lim & Normizan (2002). This data consists of questionnaire survey of 457 Universiti Utara Malaysia (UUM) graduates that attended the UUM Convocation on 14 Sept 2002. The earliest unemployment spells begins on 06 October 2001 and the censored period is fixed at 14 September 2002. In the data, 46.27% of them are classified as "unemployed" and 53.73% are classified as "employed".
The data used in this study are mainly for illustrative purpose. Indeed, the method illustrated can be easily applied into other large and more representative data set. Thus, the representativeness of the data in this study draws less important.

2.2 Methodology

The piecewise exponential proportional hazards model (flexible baseline hazard – to accommodate possible unobserved heterogeneity) is used to develop the statistical prediction

model. In addition, the Weibull model is estimated. For out-of-sample evaluation purpose, 40 observations are randomly selected out from the data. The 40 selected observations consists of 47% employed and 53% unemployed

3. Results

Table 1 presents the estimated result of the piecewise exponential model (Model I) and the Weibull model (Model II). Following Copas (1996), the statistical prediction model can be developed. The coefficients estimated (those significant at 10% level or below), are used to predict the hazard rate, i.e., the probability of exit from unemployment. Since the main purpose of this paper is to illustrate the prediction of the employability, the discussions on the significant of the variables are skipped.

Let us pretend that the 40 omitted observations are the going-to-be graduates. To identify their employability, we use the Model I to calculate their individual predicted hazard rate:

$$\hat{h}_i(t) = h_o(t)\exp(x_i\hat{\beta} + z_t\hat{\alpha}) = \exp(\hat{a})\exp(x_i\hat{\beta} + z_t\hat{\alpha}) = \exp(\hat{a} + z_t\hat{\alpha})\exp(x_i\hat{\beta})$$

where
x = the time invariant independent variables
z = the monthly dummy variables

This predicted hazard can be used as employability indicator (higher hazard rate implies higher employability) and use to rank their relative employability position compare to others. Compared to using a fixed value (for instance, hazard rate of 10% and below) for early identification, the use of relative measurement is preferable. Regardless of the employability level, some graduates have to be unemployed due to the demand constraints. Then, relatively, those rank at the bottom of employability are those identified as the high risk of unemployed.

Table 1: The estimated parametric proportional hazard model

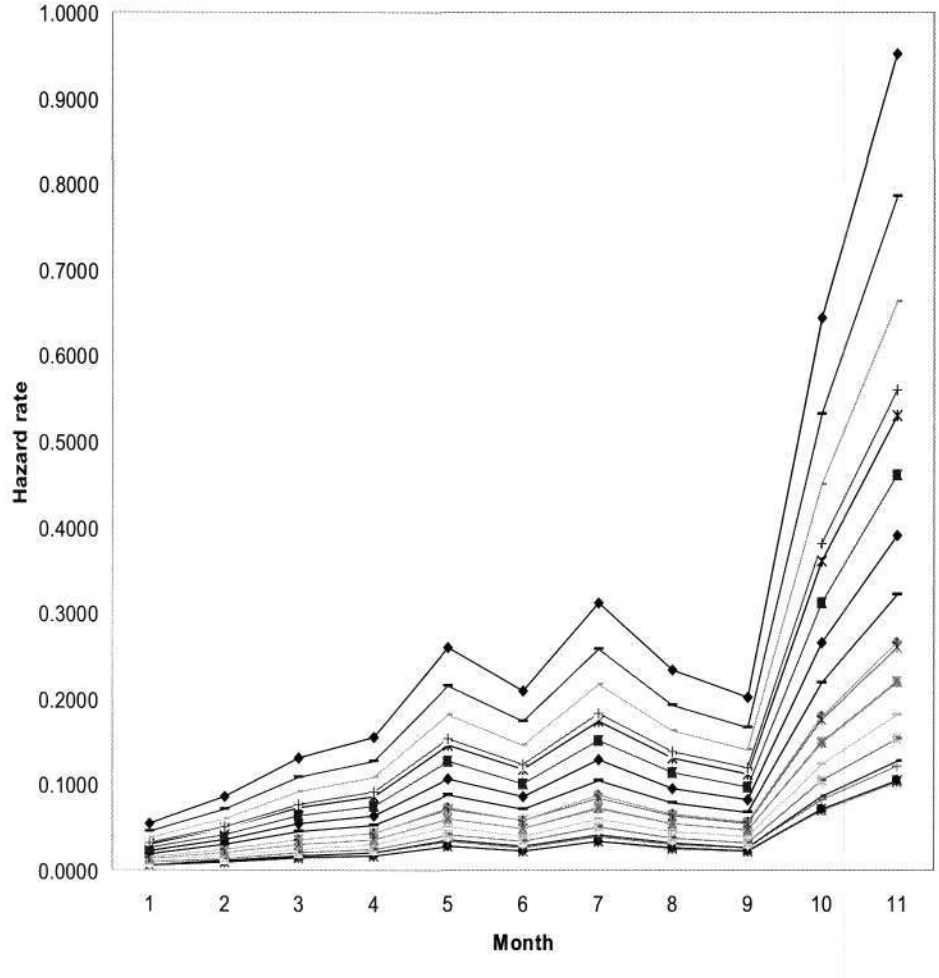
	Model I (Piecewise exponential)		Model II (Weibull)	
	Coefficient	Robust S.E	Coefficient	Robust S.E
age	-0.1915**	0.0973	-0.1570*	0.0925
marl	-0.3370	0.8082	-0.2729	0.7931
ingfstl	-0.2527	0.7015	-0.2368	0.6951
cgpa	-0.0165	0.3835	-0.0446	0.3752
uei	0.0203	0.0185	0.0188	0.0184
preuei	0.0485	0.0356	0.0471	0.0346
prakt	-0.5287**	0.2551	-0.5194**	0.2489
partt	0.3606**	0.1636	0.3477**	0.1596
Dethn1	1.3209***	0.3257	1.3044***	0.3180
Dethn2	0.7570	0.5403	0.6915	0.5340
Dethn3	0.5457	0.8161	0.4807	0.8166
Dprog1	-0.4123	0.2761	-0.4030	0.2732
Dprog2	0.7412***	0.1953	0.7298***	0.1918
Dmonth2	0.4628	0.3462		
Dmonth3	0.8788***	0.3253		
Dmonth4	1.0497***	0.3319		
Dmonth5	1.5694***	0.3184		
Dmonth6	1.3530***	0.3666		
Dmonth7	1.7487***	0.4800		
Dmonth8	1.4596***	0.5891		
Dmonth9	1.3113**	0.6322		
Dmonth10	2.4758***	0.4964		
Dmonth11	2.8651***	0.4664		
_cons	-2.5713	2.5685	-5.7007	2.4874**
p			1.6745	0.1226

Notes:
1. ***, **, and * represent significant at 1%, 5%, and 10% level respectively
2. Please refer to Appendix A for definition and measurement of the variables

Figure 1 presents the predicted hazard. The hazard rates of the going-to-be-graduates are almost similar (indeed, the predicted hazard rates are below 50% for all the 40 going-to-be-graduates) on the first nine months of their job search. This suggests that for the first nine months of job search, all the graduates are having almost similar probability of exit from unemployment, regardless their true employability level. Thus, this implies that nine months is the reasonable expected unemployment duration of in Malaysia.

From 9th month and onwards, we can clearly distinguish high and low employability students and the hazard rate is above 50% for some of the graduates. Relatively, these graduate can be identified as high employability graduates. For instance, from Figure 1, the student (id=112) has the highest employability and the student (id=49) has the lowest employability. Thus, this student (id=49) is identify as low employability. Help (or warning) can be rendered to this student. In the similar pattern, other low employability students can be identified. Thus, early identification can be implemented easily.

Figure 1: The predicted hazard rate



Since we know the employment outcomes of these 40 "pretend" going-to-be graduates, the performance of the statistical prediction model can be evaluated. Table 2 presents the result. For model I, six students are on the bottom of 15% (rank on their predicted hazard), five of them are

unemployed (correctly predicted) and one is employed (incorrectly predicted). For bottom 20%, 6 out of 8 persons are correctly predicted. For bottom 25%, 8 out of 10 are correctly predicted.

Table 2: Number correctly predicted

	Bottom		
	15% (6 persons)	20% (8 persons)	25% (10 persons)
Model I	5	6	8
Model II	5	6	7

The prediction of Weibull model (model II) has almost similar performance with piecewise exponential model (model II), except at bottom 25%, it correctly identifies 7 out of 10. This number correctly predicted (the predictive power) also can help to choose the best statistical prediction model (selection of estimation method and variables). In this case, from Table 2, the model I outperforms model II in the prediction of bottom 25%.

4. Conclusions

In conclusion, this paper illustrates that the statistical prediction model can be used for early identification of low employability graduate. The relative predicted hazard rate can serve as signaling of student's employability; and also as a tool to allocate the limited places in re-training program for unemployed graduates. The study suggests that nine months is the reasonable expected unemployment duration of graduates in Malaysia. Nevertheless, this finding is needed to be confirmed by other studies with more representative sample. Indeed, the statistical prediction model can be generalized to other estimation techniques and more representative data. It is hope that this paper will open a widely discussion on this early identification of low employability graduate especially in Malaysia.

Acknowledgement

I would like to thank Prof Dr Mohd Sahar Sawiran and the anonymous referee of Jurnal Tekonologi Maklumat & Sains Kuantitatif whose comments have improved this article. However, I am alone responsible for any remaining error that it might exist.

References

Black, A.D.,Smith, A.J.,Plesca, M.,Shannon, S. 2003. Profiling UI claimants *to allocate reemployment services: evidence and recommendations for States*, US Department of Labor ETA Occasional Papers.

Carpenter, R.G. 1983. Scoring to provide risk-related primary health care: evaluation and updating during use, Journal of the Royal Statistical Society Series C, 42, 315-331.

Copas, J. 1996. Predicting reoffending for discretionary conditional release, Research Study 150, UK Home Office.

Copas, J., Marshall, P. 1998. The offender group reconviction scale: a statistical reconviction score for use by probation officers, Journal of the Royal Statistical Society Series C , 47, 159-171.

Dabalen Andrew, Oni Bankole, Adekola A. Olatunde. 2001. Labor market prospects for university graduates in Nigeria, Higher Education Policy , 14, 141-159.

Lim, Hock-Eam & Normizan Bakar. 2002. The unemployment duration of Universiti Utara Malaysia Graduates: the impact of English language proficiency, Unpublished manuscript, Faculty of Economics, Universiti Utara Malaysia.

Morshidi Sirat, Rosni Bakar, Lim, Hock-Eam & Mohamed Nasser Katib. 2004. Pencapaian akademik & kebolehgunaan tenaga siswazah institusi pengajian tinggi, USM IPPTN Monograf 3/2004, Universiti Sains Malaysia.

OECD, 1998. Early identification of jobseekers at risk of long-term unemployment: the role of profiling. OECD, Paris.

Payne, C., Payne, J. 2000. Early identification of the long-term unemployed, UK PSI Research Discussion Paper 4.

Roed, K., Zhang, T. 2003. Does unemployment compensation affect unemployment duration? The Economic Journal, 113, 190-206.

Appendix A:
Definition and Measurement of Variables

Variable Abbreviation	Definition	Measurement
age	Age	in years
mar1	Dummy variable for marital status	1 if married 0 if single
Ingfst1	Use of English language as first language	1 if yes 0 if no
cgpa	Cumulative Grade Point Average (Academic attainment)	In continuous scale
uei	Undergraduate English language proficiency score	In continuous scale
preuei	Pre-University English language proficiency score	In continuous scale
prakt	Dummy variable for attending the industrial training	1 if yes 0 if no
partt	Dummy variable for having part-time work experience	1 if yes 0 if no
Dethn1	Dummy variable for ethnicity: Chinese (Base or comparison group =Malay)	1 if Chinese 0 if otherwise
Dethn2	Dummy variable for ethnicity: India (Base or comparison group = Malay)	1 if Indian 0 if otherwise
Dethn3	Dummy variable for ethnicity: Others (Base or comparison group = Malay)	1 if others 0 if otherwise
D2prog1	Dummy variable for degree obtained : B.Information Technology (Base or comparison group = Economics & Management)	1 if B. Information Technology 0 if otherwise
D2prog2	Dummy variable for degree obtained : B.Accouting (Base or comparison group = Economics & Management)	1 if B. Accounting 0 if otherwise
Dmonth2-11	Dummy variables for month (Base or comparison group = 1st month)	