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

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

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

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

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

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

Research Performance Evaluation using Data Envelopment Analysis (DEA)
Norshahida Shaadaw
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 Hec Tang
Prof. Madya Dalialah Abd. Ghani
Mohd Hanafi Tunin ASA

Penyunting Luaran
Prof. Madya Dr. Khairuddin Omar
Jabatan Sains dan Pengurusan Sistem
Fakulti Teknologi dan Sains Maklumat, UKM Bangi
E-mail: kjts@fms.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 Environmental Studies,
UPM Serdang
E-mail: habshah@fsas.upm.edu.my

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

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 mendedahkan pandangan dan pendirian rasmi jurnal ini.

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

Tel 03-55433329
Faks 03-55435501
e-mail mshahar@tmsk.uitm.edu.my


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
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:

$$\hat{h}(t) = h_0(t) \exp(x_1\beta + z_1\alpha) = \exp(\alpha) \exp(x_1\beta + z_1\alpha) = \exp(\alpha + z_1\alpha) \exp(x_1\beta)$$

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

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model I (Piecewise exponential)</th>
<th>Model II (Weibull)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Robust S.E</td>
</tr>
<tr>
<td>age</td>
<td>-0.1915**</td>
<td>0.0973</td>
</tr>
<tr>
<td>mar1</td>
<td>-0.3370</td>
<td>0.8082</td>
</tr>
<tr>
<td>ingfst1</td>
<td>-0.2527</td>
<td>0.7015</td>
</tr>
<tr>
<td>cgpa</td>
<td>-0.0165</td>
<td>0.3835</td>
</tr>
<tr>
<td>uei</td>
<td>0.0203</td>
<td>0.0185</td>
</tr>
<tr>
<td>preuei</td>
<td>0.0485</td>
<td>0.0356</td>
</tr>
<tr>
<td>prakt</td>
<td>-0.5287***</td>
<td>0.2551</td>
</tr>
<tr>
<td>partt</td>
<td>0.3606**</td>
<td>0.1636</td>
</tr>
<tr>
<td>Dethn1</td>
<td>1.3209***</td>
<td>0.3257</td>
</tr>
<tr>
<td>Dethn2</td>
<td>0.7570</td>
<td>0.5403</td>
</tr>
<tr>
<td>Dethn3</td>
<td>0.5457</td>
<td>0.8161</td>
</tr>
<tr>
<td>Dprog1</td>
<td>-0.4123</td>
<td>0.2761</td>
</tr>
<tr>
<td>Dprog2</td>
<td>0.7412***</td>
<td>0.1953</td>
</tr>
<tr>
<td>Dmonth2</td>
<td>0.4628</td>
<td>0.3462</td>
</tr>
<tr>
<td>Dmonth3</td>
<td>0.8788***</td>
<td>0.3253</td>
</tr>
<tr>
<td>Dmonth4</td>
<td>1.0497***</td>
<td>0.3319</td>
</tr>
<tr>
<td>Dmonth5</td>
<td>1.5694***</td>
<td>0.3184</td>
</tr>
<tr>
<td>Dmonth6</td>
<td>1.3530***</td>
<td>0.3666</td>
</tr>
<tr>
<td>Dmonth7</td>
<td>1.7487***</td>
<td>0.4800</td>
</tr>
<tr>
<td>Dmonth8</td>
<td>1.4596***</td>
<td>0.5891</td>
</tr>
<tr>
<td>Dmonth9</td>
<td>1.3113**</td>
<td>0.6322</td>
</tr>
<tr>
<td>Dmonth10</td>
<td>2.4758***</td>
<td>0.4964</td>
</tr>
<tr>
<td>Dmonth11</td>
<td>2.8651***</td>
<td>0.4664</td>
</tr>
<tr>
<td>cons</td>
<td>-2.5713</td>
<td>2.5685</td>
</tr>
<tr>
<td>p</td>
<td></td>
<td>1.6745</td>
</tr>
</tbody>
</table>

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 1, 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

<table>
<thead>
<tr>
<th></th>
<th>Bottom</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15%</td>
<td>20%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>(6 persons)</td>
<td>(8 persons)</td>
<td>(10 persons)</td>
</tr>
<tr>
<td>Model I</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Model II</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

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


### Definition and Measurement of Variables

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