Educational Research and Review Vol. 4 (5), pp. 267-271, May 2009
Available online at http://www.academicjournals.org/ERR
ISSN 1990-3839 © 2009 Academic Journals

Full Length Research Paper

Accelerated failure-time models of graduation

Justin R. Chimka* and Qilu Wang

Department of Industrial Engineering, University of Arkansas,4207 Bell Engineering Center, Fayetteville, AR 72701. USA.

Accepted 14 May, 2009

This third article in a series describing survival analysis of engineering student retention and graduation introduces accelerated failure-time as an alternative to the Cox proportional hazards model to the context of student data. The new survival analysis of graduation data presented here assumes different distributions including exponential, lognormal and Weibull, and assesses efficiency and goodness of fit based on estimated parameters, likelihood and number of observations. Results are associated with the effects of American College Test and Scholastic Assessment Test scores, gender, and other demographic information on retention and graduation. Some results confirm what we have previously learned from proportional hazards models of graduation, and some results are unique to accelerated failure-time models.

Key words: Graduation, accelerated failure-time, retention, survival analysis.

INTRODUCTION

Two frequently used models for survival analysis are the Cox proportional hazards model and the accelerated failure-time (ATF) model. We have previously used Cox proportional hazards models to study effects such as standardized test scores and gender on variation in student graduation. Proportional hazards models of graduation were originally based on main effects models of graduation, controlling for descriptors such as in-state residence, hometown population and student major (Chimka et al., 2007,2008), and later we examined interaction between pairs of previously considered main effects (Chimka and Lowe, 2008). This third article introduces the use of alternative parametric AFT models to analyze the same graduation data as before, and graduation data in general.

In our first review of the literature (Chimka et al., 2007, 2008), we found examples of modeling research into college student behavior such as logistic regression (Besterfield et al., 1997; Bruggink and Gambhir, 1996; Stage, 1988), least squares regression (Cabrera et al., 1992), event history modeling (DesJardins et al., 2002), And multilevel logit models (Smyth and McArdle, 2004), used to study student retention and graduation. Other literature that focuses on university athletic programs (Ferris et al., 2004; Mangold et al., 2003) and specific ethnic groups (Yeh, 2004,2005; Zurita, 2004, 2005) has also been reviewed. In the second article (Chimka and Lowe, 2008), more about logistic regression models of college student behavior was added (Berkovitz et al., 2006,2007; Davidson et al., 2006,2007; Scott et al., 2006), and some conclusions about factors and their effects on student retention and graduation were relayed (Gansemer-Topf et al., 2006; Randolph et al., 2006).

Recently Wohlgemuth et al. (2006, 2007) used logistic regression to estimate the likelihood of student retention for each of four years, as well as the outcome of graduation for the next three years, intending to provide insights into the influences of demographic, financial, environmental and academic characteristics. Herzog (2006) used decision trees and neural networks with a multinomial logistic regression model to predict outcomes of student retention and time to degree completion. The author showed out-performance of data-mining methods compared to traditional statistical methods. Calcagno et al. (2006) examined longitudinal transcript data to study the educational outcomes of community college students with different ages. A discrete-time hazard model was used to compare the performance of older students who

*Corresponding author. E-mail : jchimka@uark.edu. Tel.: (479) 575-7392.

Table 1. Explanatory variables.

Variable	Value
fem	1 if the student is female, 0 if the student is male
ok	1 if the student attended high school in Oklahoma, 0 otherwise
oohu	Percent owner-occupied housing units in the 3-digit ZIP Code tabulation area of the student's high school according to Census 2000
tp	Total population in the 3-digit ZIP Code tabulation area of the student's high school according to Census 2000
engact	English ACT score
mathact	Math ACT score
readact	Reading ACT score
sciact	Science ACT score
mathsat	SAT Math score
verbsat	SAT Verbal score
Engmaj	1 if the student's major is engineering, 0 otherwise (time-varying)

Table 2. Descriptive statistics.

	ACT Scores Only	SAT Scores Only	ACT Scores	SAT Scores	Entire Sample
Students	181	62	367	248	429
Grad. Rate	0.425	0.581	0.493	0.565	0.506
Proportion *fem*	0.249	0.226	0.218	0.198	0.219
Proportion *ok*	0.845	0.097	0.668	0.395	0.585
Mean *engact*	23.9	-	25.2	26.5	-
Mean *mathact*	25.4	-	26.6	27.9	-
Mean *readact*	25.4	-	27.1	28.7	-
Mean *sciact*	24.8	-	26.0	27.3	-
Mean *mathsat*	-	631	631	631	-
Mean *verbsat*	-	603	612	610	-

first entered college at age 25 or later with that of more traditional-age students.

Some articles study the effects of education programs or actions to promote student retention and success. Noble et al. (2007,2008) measured the effects of a program called ESSENCE on GPAs and graduation rates of first year students. Tinto (2006,2007) suggested some research areas that need to be explored on student retention that are more concerned with institutional action and programs. Also the author proposed that more work should be done with respect to low-income students.

The AFT model described in this article has been chosen for analysis by others in many contexts including studies of sociology (Yamaguchi, 1994; Yamaguchi, 2003), bladder cancer (Lin et al., 1998), melanoma (Cho and Schenker, 1999), AIDS (Betensky et al., 2001), cardiovascular diseases (Menotti et al., 2003) and cirrhosis (Park and Wei, 2003). However our research is believed to be the first application of AFT to study student retention and graduation.

DATA AND METHODS

The survival data used for this series of articles are observations of engineering student graduation. The cohort was followed for six and a half years, and it is composed of 429 first-time students having declared the engineering major, admitted to the University of Oklahoma in fall 1995. Fixed independent variables for all students are gender, Oklahoma residence, owner-occupied housing percentage, population of hometown, and American College Test (ACT) and Scholastic Assessment Test (SAT) scores. In Table 1 is a list of these explanatory variables along with their possible values. Table 2 includes descriptive statistics including graduation rates. For example the graduation rate among our 429 students was 0.506. Approximately 22% of them were female, and 58% were from the state of Oklahoma. Column headings of Table 2 describe three subgroups of students: Some students took ACT only, some took SAT only, and some students took ACT and SAT.

The survival time to graduation t is expressed with the natural logarithm in a linear function $\log(t_j) = x_j\beta + z_j$ where $x = (x_1, \dots x_k)$ is a vector of explanatory variables, $\beta = (\beta_1, \dots \beta_k)$ is a vector of regression coefficients reflecting the effects of the explanatory variables on survival, and $z = (z_1, \dots z_k)$ is the error with a common but completely unspecified distribution function. Survival data typi-

Table 3. AIC values.

	SAT Only	ACT Only	SAT&ACT
Lognormal	-0.02404	0.064093	0.025963
Exponential	0.83222	0.659702	0.814659
Weibull	-0.06448	0.112976	0.098552

Table 4. SAT Scores only (n = 130 observations for 62 students).

	Hazard Ratio	Std. Err.	***P-value*** **>** **|*z*|**
engmaj	1.31	0.55	0.530
fem	4.21	2.02	0.003
ok	0.79	0.64	0.767
oohu	0.97	0.03	0.370
tp	1.00	4.40e-07	0.004
mathsat	1.01	0.00	0.005
verbsat	1.00	0.00	0.151
/ln_p	2.16	0.13	0.000

ally have at least some censored observations. For instance in our models, the observations are right censored, since some students were not followed until graduation due to a time restriction in the study. The distributional form of the error z determines the regression model, which means that for each distribution of z there is a corresponding distribution of t. It should be noted that all AFT models are named for the distribution of t. If the distribution of z is extreme-value density with one or two parameters, the exponential and the Weibull distribution of t can be obtained respectively. Similarly, if the distribution of z is set to be logistic or normal, thedistribution of t will be log-logistic or log-normal respectively. And the log-gamma density yields a gamma regression model. In our case, three different AFT models were estimated (log-normal, exponential and Weibull), and the results were compared using the Akaike Information Criterion (AIC).

The vector of regression coefficients is chosen to be the maximum (log) likelihood vector. For example an expression of (natural) log likelihood for the gamma distribution family is $\log L = -y / \mu + \log(1/\mu)$, where y and μ indicate observations and expectations, respectively. Finding the maximum likelihood model associated with relevant observations is a nonlinear optimization problem solved with a computer. Estimation for research described here was done with Stata Statistical Software (StataCorp, 2007).

Akaike (1974) proposed a method to systematically handle statistical model selection. AIC penalizes the log likelihood of each particular model to reflect the number of parameters being estimated and observations used to fit the model. AIC can be defined as AIC = $(2/n)$ (k – log likelihood), where n is the number of observations, and k is the number of estimated parameters which includes the number of model covariates, a constant, and shape parameter (if necessary). The model with the lesser value of AIC is considered to be the better model.

Statistical models

We began by analyzing three datasets described by student choice of standardized test:

1) ACT score only.
2) SAT score only.
3) Both ACT and SAT, with log-normal, exponential and Weibull regression models for each.

These are nine models in all. The common variables for the nine are engmajor, fem, ok, oohu and tp (Table 1). For the model of ACT only, extra variables that related to ACT including engact, mathact, readact and sciact were added. For the model of SAT only, mathsat and verbsat were added. All these variables were considered in the models of both ACT and SAT. One model was selected for each of the first two student groups by computing and comparing AIC values (Table 3). Among models of students with only SAT scores, Weibull has the least AIC. Among models of students with only ACT scores, lognormal has the least AIC. We found not one model was statistically significant as a whole among those of students with both ACT and SAT scores.

For students submitting only SAT scores upon application (n = 130 observations for 62 students), results indicate the model as a whole is statistically significant (global P-value > χ^2 = 0.0000), but since some continuous main effects are not significant we analyzed relevant interactions. We considered the interaction between oohu and verbsat, but it turned out the interaction is not significant (P-value >│z│= 0.137), so the model without interaction was kept, and we conclude that female students (P-value >│z│= 0.003) and students with better SAT Math scores (P-value >│z│= 0.005) are more likely to graduate. Table 4 provides details of these results that are the same as results from the first article (Chimka et al., 2007, 2008).

For students with ACT scores only (n = 392 observations for 181 students), results indicate that the model as a whole is statistically significant (global P-value > χ^2 = 0.0209), but the individual regression coefficients are not. Consider that "if the F-test for significance of regression is significant, but tests on the individual regression coefficient are not significant, multicollinearity may be present (Montgomery and Runger, 2007)." Several remedial measures can be used to solve the problem of multicollinearity. One is to delete certain variables from the model. In our case, nine new models were estimated, each without a different, previously considered independent variable. Among these nine, we looked for the ones with significant main effects, and that yielded four models: the model with engmajor missing, the model with mathact missing, the model with readact missing, and the model with *sciact* missing. By computing the AIC values again, we found the one with engmajor missing has the least AIC value, so it was selected. This model as a whole is statistically significant (global P-value > χ^2 = 0.0169), and it can be concluded that students with better ACT Math scores (P-value >│z│= 0.033) are more likely to graduate. See Table 5 for details of this model, and remember it came at the expense of having lost the ability to control for whether or not a student has the engineering major. Again this was necessary to eliminate problems of multicollinearity.

Since we found no significant AFT model of students having taken both ACT and SAT tests, models of all students that took the SAT test and all those that took the ACT test were estimated separately. Please note the group "students that took SAT" includes two subgroups of students: one that took SAT only and one that took SAT and ACT. Likewise the group "students that took ACT" refers to one subgroup that took ACT only plus one subgroup that took SAT and ACT.

Between models of all students that took ACT we selected the lognormal one with greatest value of log likelihood (Table 6). The model as a whole is statistically significant (global P-value > χ^2 = 0.0169). The results show that students with better ACT Science

Table 5. ACT scores only (n =392 observations for 181 students).

	Coefficient	Std. Err.	P-value > |z|
fem	-0.12	0.04	0.789
ok	0.06	0.05	0.226
oohu	0.00	0.00	0.252
tp	1.00E-07	6.70E-08	0.135
engact	0.00	0.01	0.802
mathact	-0.01	0.01	0.033
readact	0.01	0.01	0.239
sciact	-0.01	0.01	0.132
/ln_sig	-1.71	0.08	0.000

Table 6. ACT scores (n = 777 observations for 367 students).

	Coefficient	Std. Err.	P-value > |z|
engmajor	-0.01	0.03	0.836
fem	0.00	0.03	0.906
ok	0.03	0.03	0.219
oohu	0.00	0.00	0.097
tp	5.25E-08	3.00E-08	0.080
engact	-0.00	0.00	0.873
mathact	-0.01	0.00	0.100
readact	0.00	0.00	0.233
sciact	-0.01	0.00	0.037
sat	-0.03	0.03	0.312
/ln_sig	-1.72	0.05	0.000

Table 7. SAT scores (n = 515 observations for 248 students)

	Coefficient	Std. Err.	***P-value*** > |***z***|
engmaj	0.13	0.03	0.624
fem	-0.02	0.03	0.639
ok	0.03	0.03	0.332
oohu	0.00	0.00	0.263
tp	7.11e-08	2.63e-08	0.007
mathsat	-0.00	0.00	0.047
verbsat	0.00	0.00	0.881
act	-0.01	0.03	0.874
/ln_sig	-1.76	0.06	0.000

scores (*P*-value >$|z|$= 0.037) are more likely to graduate.

Between models of all students that took the SAT we again selected the lognormal one, and we were interested in whether or not some insignificant main effects are important. Therefore we assessed the interaction between oohu and verbsat. After studying the new model with interaction, we found interaction is not significant (P-value >$|z|$ = 0.638). Therefore, we reverted to the original model that controls for oohu and verbsat (Table 7). The model as a whole is statistically significant (global *P*-value > χ^2 = 0.0080). It indicates that students with better SAT Math scores (*P*-value >$|z|$ = 0.047) are more likely to graduate. However, those whose high school area has larger population (*P*-value >$|z|$ = 0.007) are less likely to graduate. This is an interesting result typical only for students who took the SAT test. It may give us some insight into the relationship between geographic effects and student graduation rate, because the SAT test and ACT test are geographically popular within different areas of the US.

DISCUSSION

Results that are consistent with previous research into proportional hazards models of graduation are that female students and students with better SAT Math scores were more likely to graduate among those who submit only SAT scores upon application. Some new results include the following. For those who submitted only ACT scores upon application, students with better ACT Math scores were more likely to graduate. For the group of all students who submitted ACT scores, students with better ACT Science scores were more likely to graduate. For all those who submitted SAT scores, students with better SAT Math scores were more likely to graduate. In this group also those whose high school area had greater population were less likely to graduate.
From the results of descriptive statistics (Table 2), the mean ACT scores of students who took both ACT and SAT are higher than those who took ACT only. Note that similar results also hold true for the SAT scores, indicating those who took ACT and SAT both may be more competitive and motivated than those students having taken either one or the other of the standardized tests.

Perhaps even more can be learned about the standardized test taking decision and its relationship with retention and graduation by observing relevant graduation rates. For example the least graduation rate (0.425) is that among students with ACT scores only. Next is the rate among all students with ACT scores (0.493). Added improvement is found among all students with SAT scores (0.565). And the greatest graduation rate (0.581) in this comparison is among students with SAT scores only.

If interest in AFT models of graduation continues, then future work should concern better understanding of geographic and demographic effects on graduation rate, especially since the population of the tabulation area of the student's high school has shown to be an important factor. Such more focused investigations into demogra phic effects should also address differences we have observed across choice of standardized tests. Finally demographic effects on retention and graduation would ideally be studied with standardized test taking behavior and scores simultaneously because of geographic preferences for ACT and SAT.

REFERENCES

Akaike H (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19 (6), 716–723.

Berkovitz RA, O'Quin K (2006-2007). Predictors of graduation of readmitted "at risk" college students. J. Coll. Stud. Retent., 8(2): 199-214.

Besterfield SME, Atman CJ, Shuman LJ (1997). Characteristics of freshman engineering students: models for determining student attrition and success in engineering. J. Eng. Educ., 86(2): 139-149.

Betensky RA, Rabinowitz D, Tsiatis AA (2001). Computationally simple accelerated failure time regression for interval censored data. Biometrika. 88(3): 703-711.

Bruggink TH, Gambhir V (1996). Statistical models for college admission and enrollment: a case study for a selective liberal arts college. Res. Higher Educ. 37(2).

Cabrera AF, Nora A, Casteneda MB (1992). The role of finances in the persistence process: a structural model. Res. Higher Educ., 33(5).

Calcagno JC, Crosta P, Bakley T, Jenkins D (2006). Stepping stones to a degree: The impact of enrollment pathways and milestones on community college student outcomes. CCRC Working Paper, No.4

Chimka JR, Reed-Rhoads T, Barker K (2007-2008) Proportional hazards models of graduation. J. Coll. Stud. Retent., 9(2): 221-232.

Chimka JR, Lowe LH (2008) Interaction and survival analysis of graduation data. Educ. Res. Rev., 3(1):29-32

Cho M, Schenker N (1999). Fitting the log-F accelerated failure time model with incomplete covariate data. Biometrics, 55, 826-833.

Davidson WB, Beck HP (2006-2007). Survey of academic orientations scores and persistence in college freshmen. J. Coll. Stud. Retent., 8(3): 297-305.

DesJardins SL, McCall BP, Ahlburg DA, Moye MJ (2002). Adding a timing light to the 'tool box'. Res. Higher Educ. 43(1).

Ferris E, Finster M, McDonald D (2004). Academic fit of student-athletes: an analysis of NCAA division I-A graduation rates. Res. Higher Educ., 45(6).

Herzog S (2006). Estimating student retention and degree-completion time: decision trees and neural networks Vis-à-Vis regression. New Directions for institutional Research, no. 131, fall 2006.

Lin DY, Wei LJ, Ying Z (1998). Accelerated failure time models for counting processes. Biometrika, 85(3): 605-618.

Mangold WD, Bean L, Adams D (2003). The impact of intercollegiate athletics on graduation reates among major NCAA division I universities. J. Higher Educ., 74(5).

Menotti A, Puddu PE, Lanni M, Kromhou D, Blackburn H, Nissinen A (2003). Twenty-five-year coronary mortality trends in the seven countries stud using the accelerated failure time model. Eur. J. Epidemiol., 18, 113-122.

Montgomery DC, Runger GC (2007). Applied statistics and probability for engineers (4^{th} ed.), John Wiley & Sons, Inc.

Noble K, Flynn NT, Lee JD, Hilton D (2007-2008). Predicting successful college experiences: Evidence from a first year retention program. J. coll. Stud. Retent., 9(1): 39-60.

Park Y, Wei LJ (2003). Estimating subject-specific survival functions under the accelerated failure time model. Biometrika, 90(3): 717-723.

Randolph KA, Fraser MW, Orthner DK (2006). A strategy for assessing the impact of time-varying family risk factors on high school dropout. J. Family Issues, 27(7): 933-950.

Scott M, Bailey T, Kienzl G (2006). Relative success? Determinants of college graduation rates in public and private colleges in the US. Res. Higher Educ., 47(3): 249-279.

Smyth FL, McArdle JJ (2004). Ethnic and gender differences in science graduation at selective colleges with implications for admission policy and college choice. Res. Higher Educ., 45(4).

Stage FK (1988). University attrition: LISREL with logistic regression for the persistence criterion. Res. Higher Educ., 29(4).

StataCorp (2007). Stata Statistical Software: Release 10. College Station, TX: StataCorp LP.

Tinto V (2006-2007). Research and practice of student retention: what next? J. coll. Stud. Retent. 8(1): 1-19.

Wohlgemuth D, Whalen D, Sullivan J, Nading C, Shelley M, Wang Y (2006-2007). Financial, academic, and environmental influences on the retention and graduation of students. J. Coll. Stud. Retent., 8(4): 457-475.

Yamaguchi K (1994). Some accelerated failure-time regression models derived from diffusion process models. Sociol. Methodol., 24: 267-300.

Yamaguchi K (2003). Accelerated failure-time mover-stayer regression models for the analysis of last-episode data. Sociol. Methodol., 33: 81-110.

Yeh TL (2004-2005). Issues of college persistence between Asian and Asian Pacific American students. J. Coll. Stud. Retent., 6(1): 81-96.

Zurita M (2004-2005). Stopping out and persisting: experiences of Latino undergraduates. J. Coll. Stud. Retent., 6(3): 301-324.