22nd Nordic Conference on Mathematical Statistics (NORDSTAT 2008)

Analyzing dropout rates of B.Sc students

L. Camilleri* and S.R. Caruana
Department of Statistics and Operations Research, University of Malta, Msida, Malta

Summary. *Survival analysis is a useful statistical technique for analyzing failure time data. It overcomes the limitations of cross-sectional analysis and convention regression analysis. This study proposes three statistical models that analyze the failure time when university students withdraw from B.Sc courses. The sample comprises all the 91 students who commenced their four-year studies with the Faculty of Science in 2002. The first approach uses the Kaplan-Meier product limit method for estimating the survival functions under non-informative censoring. The second approach uses the Cox regression model, which involves the assumption of proportional hazard functions. The third approach uses a parametric model to estimate the hazard function using an appropriate distribution. The aim of the study is to fit survival models that predict the probabilities of retention and dropout of B.Sc students within each study area using the facilities of SPSS and STATA.*

Keywords: *Kaplan Meier estimate, Cox regression, Censoring, Lognormal distribution, Maximum likelihood estimation*

Received:
Published online:

1. Introduction

Survival analysis is a useful statistical technique for answering questions that deal with the duration of events and was originally developed by biostatisticians to model human lifetimes. The term survival data has been used for data involving time to a certain event such as relapse, death, and onset of a disease. Recently, applications of survival analysis have been extended beyond biomedical research to other fields such as criminology, sociology, marketing, health insurance practice and institutional research. The key feature of survival data is censoring. This occurs when the value of an observation is partially known. The mechanisms that give rise to censoring play an important role when making statistical inference. The purpose of this study is to conduct survival analysis to investigate the duration before withdrawal of students from B.Sc courses in the Faculty of Science. Investigation into retention and dropout is crucial to academic planning.

The data set comprises 91 B.Sc students who commenced their studies with the Faculty of Science in 2002. The course spans over four years; however, a student who fails to satisfy the requirements to proceed to the subsequent year is allowed to repeat the study programme for that year. This concession is approved only once and if a student fails twice, his course is terminated. So a student has to complete a B.Sc course successfully in at most five years. A small group of students, who enrolled for the B.Sc course in 2001 but failed during their first year of study, were included with the 2002 cohort on re-applying for the course. These students were left censored. All students who failed to complete the course for academic and non-academic causes were regarded uncensored observations; whereas students who completed the course successfully were right censored. Students applying for a B.Sc course have to choose two subject areas that include Biology, Mathematics, Physics, Chemistry and Statistics & Operations Research (SOR).

		Completed course after 4 years	Completed course after 5 years	Failed to complete course
Biology	Count	29	5	8
	Percentage	69.0%	11.9%	19.0%
Chemistry	Count	30	7	13
	Percentage	60.0%	14.0%	26.0%
Maths	Count	10	6	24
	Percentage	25.0%	15.0%	60.0%
Physics	Count	6	5	24
	Percentage	17.1%	14.3%	68.6%
SOR	Count	3	5	7
	Percentage	20.0%	33.3%	46.7%
Total	Count	78	28	76
	Percentage	42.9%	15.4%	41.8%

Table 1: Crosstab displaying frequency and percentage of students categorized by subject area and course outcome

Table 1 demonstrates that a large proportion of Biology and Chemistry students completed the B.Sc course in four years; whereas a large proportion of Mathematics, Physics and SOR students failed to complete the course. The chi-squared test ($\chi^2 = 41.38$ and $p < 0.0005$) reveals that the above association is significant and not attributed to chance. This discrepancy is mainly attributed to different assessment methods employed by the departments. Some subjects are assessed solely by examinations; whereas other subjects are assessed by both examinations and coursework. Contrasting examination paper settings and different marking schemes may add to this discrepancy.

Survival time depends on past academic performance. Students with low A-level grades are more likely to fail or withdraw than students with high A-level grades. To investigate this issue the A-level grades obtained by each student were converted to scores using the MATSEC point system where grades A, B, C, D and E correspond to 30, 24, 18, 12 and 6 points respectively. An entry qualification score was generated by summing the scores for the two A-levels.

	Mean Entry Score	Std. Deviation	95% Confidence Interval for Mean	
			Lower Bound	Upper Bound
Biology	38.86	7.754	36.44	41.27
Chemistry	38.40	8.221	36.06	40.74
Mathematics	37.11	10.702	33.59	40.62
Physics	38.57	10.007	35.13	42.01
Statistics	31.20	8.239	26.64	35.76

Table 2: Table displaying means, standard deviations and 95% confidence limits of entry scores categorized by subject area

Table 2 displays that the mean entry scores for Biology, Chemistry, Mathematics and Physics students are higher ($F = 2.587$ and $p < 0.0386$) than those obtained by SOR students. This difference is attributed to the fact that the entry requirements for some departments are more stringent than others.

	Mean Entry Score	Std. Deviation	95% Confidence Interval for Mean	
			Lower Bound	Upper Bound
Completed course after 4 years	41.85	9.121	39.79	43.90
Completed course after 5 years	35.57	8.153	32.41	38.73
Failed to complete course	34.05	7.827	32.24	35.87

Table 3: Table displaying means, standard deviations and 95% confidence limits of entry scores categorized by course outcome

Table 3 shows that the mean entry score for students who complete the course in 4 years are significantly higher ($F = 17.118$ and $p < 0.0005$) compared to the mean score of their counterparts. Students with a low entry score are more likely to fail or withdraw from the course than those with a high entry score.

2. Kaplan-Meier (product-limit) estimator

The proportion of B.Sc students who fail to complete the course at the end of each academic year depends on both the subject areas chosen and the entry qualification score of each student. These predictors contribute significantly in explaining variation of survival times. To measure survival probabilities for the distinct categories of these predictors, the students' entry scores were categorized into three groups. Entry scores ranging from 20 to 30 correspond to low grades. Moderately good and high grades range from 31 to 50 and 51 to 60 respectively. The objective of this study is to predict the proportion of unsuccessful students that drop out each year form these B.Sc course.

The survival function $S(t)$ provides the probability that a student does not withdraw from the course before time *t*. The Kaplan-Meier estimator is a nonparametric maximum likelihood estimate of $S(t)$. This estimator is obtained by maximizing the likelihood function which is expressed as the product of independent binomial likelihoods.

$$L = \prod_{j=1}^{k} \left(\lambda_j\right)^{d_j} \left(1-\lambda_j\right)^{n_j-d_j} \quad (1)$$

d_j and n_j are respectively the number of students who withdraw from the course and the number of students who are allowed to proceed at time t_j; whereas λ_j is the hazard at this time. The maximum likelihood estimator of this hazard is given by:

$$\hat{\lambda}_j = d_j / n_j \text{ for } j = 1, 2, ..., k \quad (2)$$

The Kaplan-Meier estimate is derived by replacing the hazard λ_j by its maximum likelihood estimate $\hat{\lambda}_j$.

$$\hat{S}(t) = \prod_{t_j \le t} \left(1-\hat{\lambda}_j\right) = \prod_{t_j \le t} \left(\frac{n_j - d_j}{n_j}\right) \quad (3)$$

Let the number of censored lives in the interval (t_j, t_{j+1}) be denoted by c_j. The relationship between c_j, n_j and d_j is given by:

$$n_j = d_j + c_j + n_{j+1} \text{ for } j = 1, 2, \cdots, k \quad (4)$$

Subject	Year	n_j	d_j	c_j	Survival Probability	St Error
Biology	1	42	6	0	0.8571	0.0540
	2	36	1	0	0.8333	0.0575
	3	36	1	0	0.8095	0.0606
	4	34	0	29	0.8095	0.0606
	5	5	0	5	0.8095	0.0606
Chemistry	1	50	9	0	0.8200	0.0543
	2	41	2	0	0.7800	0.0586
	3	39	2	0	0.7400	0.0620
	4	37	0	30	0.7400	0.0620
	5	7	0	7	0.7400	0.0620
Maths	1	40	16	0	0.6000	0.0775
	2	24	3	0	0.5250	0.0790
	3	21	3	0	0.4500	0.0787
	4	18	2	10	0.3375	0.0907
	5	6	0	6	0.3375	0.0907
Physics	1	35	16	0	0.5429	0.0842
	2	19	4	0	0.4286	0.0836
	3	15	2	0	0.3714	0.0817
	4	13	2	6	0.2653	0.0862
	5	5	0	5	0.2653	0.0862
Statistics	1	15	3	0	0.8000	0.1033
	2	12	2	0	0.6667	0.1217
	3	10	2	0	0.5333	0.1218
	4	8	0	3	0.5333	0.1218
	5	5	0	5	0.5333	0.1218

Table 4: Kaplan Meier survival probability estimates of B.Sc students categorized by subject area.

Table 4 displays the Kaplan-Meier survival probabilities for each subject area. The drop in survival probabilities is most conspicuous during the first year of study and this applies to all subject areas. Moreover, these survival probabilities vary considerably between subject areas. Biology students are most likely to complete the course successfully; whereas Physics and Mathematics students are most likely to fail or withdraw from the course.

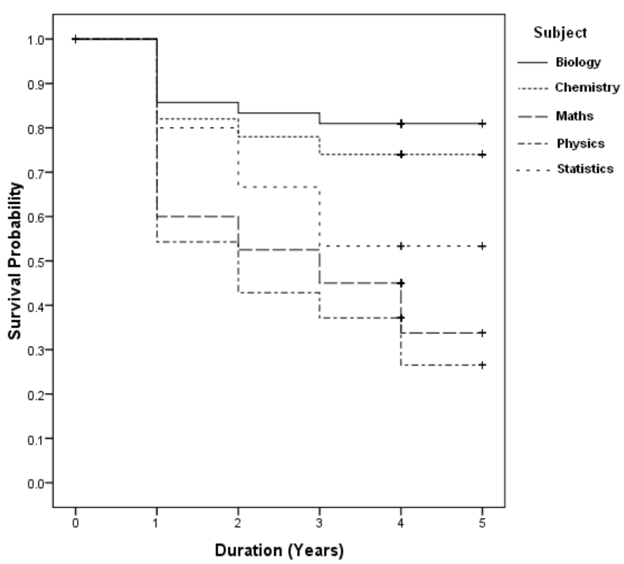

Figure 1: Survival curves of B.Sc students by subject area

The survival curves, shown in Figure 1, display a picture of the survival rate for each subject area. It is evident that most failures and drop-outs occur during the first year of study. The Log Rank (Mantel-Cox) test reveals that these survival functions differ significantly between study areas ($\chi^2 = 39.198$ and $p < 0.0005$).

Score	Year	n_j	d_j	c_j	Survival Probability	St Error
20 - 30	1	66	30	0	0.5455	0.0613
	2	36	6	0	0.4545	0.0613
	3	30	6	0	0.3636	0.0592
	4	24	0	10	0.3636	0.0592
	5	14	0	14	0.3636	0.0592
31 - 50	1	96	18	0	0.8125	0.0398
	2	78	6	0	0.7500	0.0442
	3	72	4	0	0.7083	0.0464
	4	68	4	50	0.5509	0.0782
	5	14	0	14	0.5509	0.0782

Table 5: Kaplan Meier survival probability estimates of B.Sc students categorized by entry qualification score.

The Kaplan-Meier survival probabilities, shown in Table 5, reveal that almost 45% of all students with an entry score ranging from 20 to 30 fail or drop out of the course during the first year of study. Only 36% of this low entry score group manages to complete the course successfully. The failure rate of students with middling entry score is more gradual; however, only 55% of the students in this group succeed to get a degree. The survival probabilities of the high entry score group are not displayed since all the students completed the course successfully in four years. The Log Rank (Mantel-Cox) test that compares the survival functions reveals that these survival curves differ significantly ($\chi^2 = 27.089$ and $p < 0.0005$).

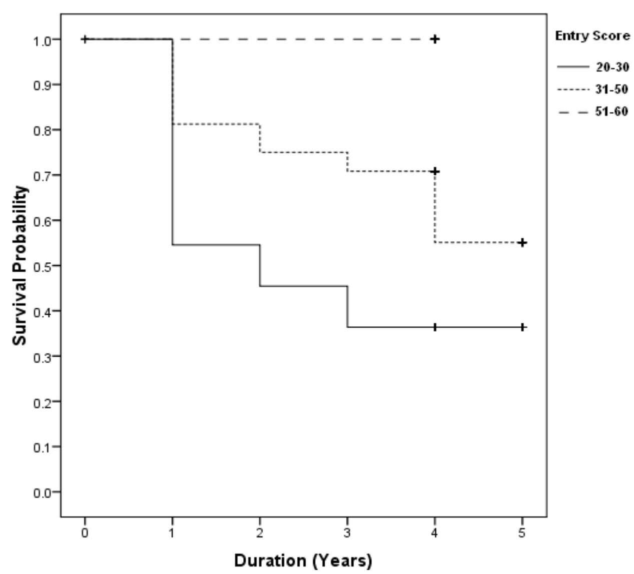

Figure 2: Survival curves of B.Sc students by entry score

3. Cox proportional hazard model

The second approach uses the Cox regression model, which involves the assumption of proportional hazard functions. This semi-parametric approach overcomes the limitations of the first approach by accommodating the effects of covariates on survival. A proportional hazards model proposed by Cox (1972) assumes that

$$h(t;\mathbf{X}) = h_o(t)\exp(\mathbf{X}'\boldsymbol{\beta}) \tag{5}$$

$h_o(t)$ is the baseline hazard function, $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)$ is a vector of regression parameters and matrix $\mathbf{X}$ includes the values of the entry qualification scores and subject area of each student.

The vector of regression parameters $\boldsymbol{\beta}$ is estimated by maximizing the partial likelihood function with respect to the parameters.

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{k} \frac{\exp(\mathbf{X}_j'\boldsymbol{\beta})}{\sum_r \exp(\mathbf{X}_r'\boldsymbol{\beta})} \tag{6}$$

The utility of the Cox model arises from the fact that the baseline hazard function determines the general shape of the hazard for all students, whereas the exponential term accounts for the differences between students. Under the Cox regression model, the hazards of two students, with predictor matrix $\mathbf{X}_1$ and $\mathbf{X}_2$, are in the same proportion at all times.

$$\frac{h(t;\mathbf{X}_1)}{h(t;\mathbf{X}_2)} = \exp\left(\mathbf{X}_1'\boldsymbol{\beta} - \mathbf{X}_2'\boldsymbol{\beta}\right) \text{ for } t \geq 0 \tag{7}$$

The model fitting process uses the likelihood ratio statistic (scaled deviance) to select the main effects and interaction terms that have a significant effect on the model fit. The scaled deviance D compares the likelihood of the current model L_c to the likelihood of the full model L_f

$$D = -2\log\left(L_c / L_f\right) \tag{8}$$

D has a chi-squared distribution with $(p-q)$ degrees of freedom, where p and q are the number of independent parameters estimated for the two models. Using a forward procedure, the parsimonious model included solely main effects.

Table 6 displays the regression coefficients, their standard errors, the Wald test statistics, and the relative hazards. Relative hazards are exponentiated regression parameters that are interpreted as the hazard change for specified risk-related predictors compared to the baseline. The relative hazard for Entry Score is 0.950, which implies that for every unit increment in the entry score the odds of failing to complete the course decreases by 5%.

Predictor	Estimate	St Error	Wald	df	P-value	Relative Hazard
Entry Score	-0.051	.014	12.55	1	0.000	0.950
Biology	-0.496	.530	0.875	1	0.350	0.609
Chemistry	-0.200	.480	0.174	1	0.677	0.819
Mathematics	0.574	.438	1.718	1	0.190	1.776
Physics	0.896	.439	4.165	1	0.041	2.449
Statistics	0					1

Table 6: Parameter estimates and Relative hazards

Parameter estimates for the subjects vary considerably. Positive parameters indicate higher risk of failure. This implies that Mathematics and Physics students are more likely to fail the B.Sc course than Biology and Chemistry students. The odds that a Mathematics student withdraws or fails to complete the course is 77.6% higher compared to a Statistics student. Conversely, the odds of failing for a Biology student is 39.1% lower compared to a Statistics student. The parameter for Statistics is set to 0 (intrinsic aliasing) to accommodate redundancy in the specification of the linear structure.

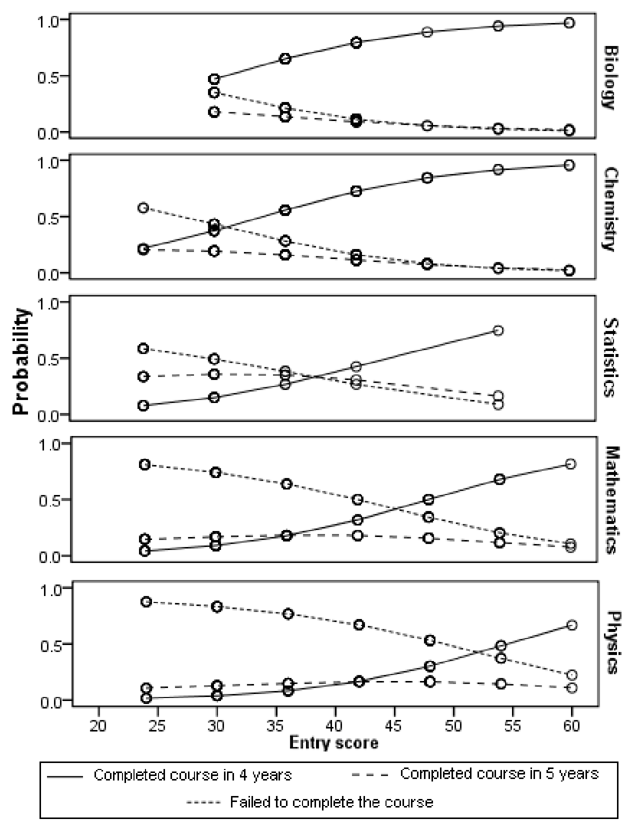

Figure 3: Probability plots categorized by course outcome, subject area and entry qualification score.

Figure 3 displays the probability curves for the course outcome categories by subject area and entry qualification score. These probability distributions are obtained by fitting a Multinomial Logistic Regression model having course outcome as the response variable and entry score and subject area as predictors. The plots reveal that for all subjects the probability of failing to complete the course decreases with an increase in the entry qualification score. The probability curves representing completion in 4 years and failure to complete course overlap at different entry scores for the five subject areas. For a Chemistry student this overlap occurs at entry score of 31; whereas, for a Physics student this overlap occurs at an entry score of 52. For Statistics and Mathematics students this overlap occurs at entry scores of 38 and 45 respectively. The probability that Biology students complete the course in 4 years exceeds the probability of failing irrespective of the entry score. This implies that Biology students with low entry scores are less likely to withdraw or fail to complete the course than Physics and Mathematics students.

4. Parametric Regression model

The Cox proportional hazard regression model is a widely used tool in analyzing censored survival data. However, constraints arise when using this model. They include the restrictive assumption of proportional hazard for covariate effects. The hazard function, which is the instantaneous, failure rate at any point in time, is essential to predict the probabilities of retention and dropout of B.Sc students. A smooth estimate of the hazard function can be obtained by implementing parametric regression models.

There are several contender models with different hazard function. The Exponential survival model is the simplest parametric model and specifies that the hazard does not vary with time. The Weibull model specifies a monotonic hazard function adequate for ageing processes where the risk of failure increases monotonically with an increase in time or for a declining hazard after a medical treatment where the risk of failure decreases monotonically as time increases. The Gompertz model specifies an exponential hazard function. This model is appropriate when hazard risk increases or decreases exponentially. The Lognormal and Log-Logistic models specify humped hazard rates, specifically initially increasing then decreasing rates.

A common approach for comparing non-nested models is to use information criteria that are based on the bias-corrected log-likelihood given by:

$$C = -2\log L(\mathbf{\Psi}) + dc \qquad (9)$$

where d is the number of estimated parameters and c is a penalty constant. The second term, which is the penalty term, measures the complexity of the model. The Akaike (AIC) and Bayesian (BIC) information criteria arise when $c = 2$ and $c = \log(N)$ respectively, where N is the sample size. When the number of parameters in a model fit is increased, the log-likelihood decreases but the penalty term increases. So the AIC and BIC criteria trade-off between complex models that explain the data well but comprise too many parameters and simpler models that explain the data adequately but have less parameters. The model with the smallest AIC or BIC value is the model that fits the data best (parsimonious model).

Distribution	Log likelihood	AIC	BIC
Exponential	-154.96	321.92	340.59
Weibull	-154.61	323.22	345.01
Gompertz	-153.25	320.49	342.28
Lognormal	-149.12	312.23	334.02
Log-logistic	-150.25	314.50	336.28

Table 7: AIC and BIC values for contender survival models

The AIC and BIC criteria both infer that the lognormal model is the most plausible. Figure 4 displays the hazard function which peaks at around the first year of study. This conforms to what we expect since a large proportion of dropouts occur after the first year of study. As expected the lognormal distribution exhibits humped hazard rates

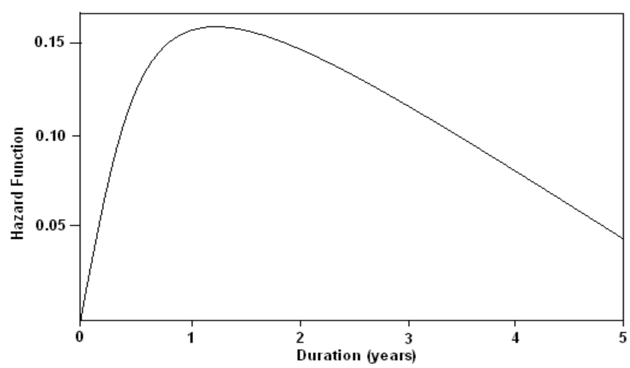

Figure 4: The Lognormal distributed hazard function

For the Lognormal model, the logarithm of time follows a Normal distribution having density function

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[\frac{-1}{2\sigma}\{\log(t) - \mu\}\right] \qquad (10)$$

The Lognormal survival function is given by:

$$S(t_j) = 1\text{-}\Phi\left[\frac{\log(t_j) - \mu_j}{\sigma}\right] \qquad (11)$$

$\Phi(z)$ is the standard normal distribution function. The lognormal regression is applied by setting $\mu_j = \mathbf{x}'_j \boldsymbol{\beta}$ and treating the standard deviation σ as a scale parameter to be estimated from the data.

Table 8 displays the parameter estimates, standard errors and their 95% confidence limits. The parameter estimate for entry score is significant implying that it is a crucial predictor of failure time. The parameter estimate (-1.198) for Physics is significantly different from 0 implying that failure times for Physics students differ considerably from those of Statistics students.

Predictor	Estimate	Standard Error	P-value	95% Confidence Interval	
Constant	0.0639	0.5717	0.911	-1.0566	1.1845
Entry Score	0.0566	0.0139	0.000	0.0294	0.0839
Biology	0.0999	0.4854	0.837	-0.8514	1.0512
Chemistry	-0.0205	0.4744	0.966	-0.9503	0.9093
Mathematics	-0.8408	0.4672	0.072	-1.7566	0.0750
Physics	-1.1979	0.4798	0.013	-2.1383	-0.2574
Statistics	0				
σ	1.1916	0.1232		0.9730	1.4594

Table 8: Parameter estimates and 95% confidence limit

Table 9 displays the probabilities of survival for B.Sc students throughout the study period given the predictors. The probability that a Physics student with an entry score of 24 completes the course in 4 years is 0.165; whereas the probability of survival for a Biology student with the same entry score is 0.545. These survival probabilities increase as the entry score augments.

Predictor	Entry Score	μ_j	$S(1)$	$S(2)$	$S(3)$	$S(4)$
Biology	24	1.5222	0.899	0.757	0.639	0.545
Chemistry		1.4018	0.880	0.724	0.600	0.505
Maths		0.5815	0.687	0.463	0.332	0.250
Physics		0.2244	0.575	0.347	0.232	0.165
Statistics		1.4223	0.884	0.730	0.607	0.512
Biology	36	2.2014	0.968	0.897	0.823	0.753
Chemistry		2.0810	0.960	0.878	0.795	0.720
Maths		1.2607	0.855	0.683	0.554	0.458
Physics		0.9036	0.776	0.570	0.435	0.343
Statistics		2.1015	0.961	0.881	0.800	0.726
Biology	48	2.8806	0.992	0.967	0.933	0.895
Chemistry		2.7602	0.990	0.959	0.918	0.876
Maths		1.9399	0.948	0.852	0.760	0.679
Physics		1.5828	0.908	0.772	0.658	0.565
Statistics		2.7807	0.990	0.960	0.921	0.879
Biology	60	3.5598	0.999	0.992	0.981	0.966
Chemistry		3.4394	0.998	0.989	0.975	0.958
Maths		2.6191	0.986	0.947	0.899	0.850
Physics		2.2620	0.971	0.906	0.836	0.769
Statistics		3.4599	0.998	0.990	0.976	0.959

Table 9: Survival Probabilities categorized by subject area, entry qualification score and duration

5. Conclusion

In the review we gave an overview of three approaches that are used to analyze survival data related to dropouts of students from B.Sc courses. All three models suggest that entry qualification score and study area selected by students are two central predictors of the duration of B.Sc students before withdrawing from their course. Biology students with a high entry score are the most likely to complete the course in four years. Physics students with a low entry score are the most likely to fail to complete the course.

A limitation of the study is that it does not discriminate between students who fail to qualify because they do not possess enough ECTS credits and students who resign because they simply lose interest or find the course too difficult. Another limitation is the implicit assumption is that there is no variability in survival probabilities beyond that which is explained by the predictors included in the model.

We suggest three approaches for future research. The first recommendation is to discriminate between students who fail and students who drop out. The second suggestion is to include demographic and student-related predictors in the model to improve prediction. The third suggestion is to develop frailty models that include random effects to explain unobserved heterogeneity in models for survival data. Frailty models are extensions of Cox proportional hazard models that assume that the proportionality factor, which modifies the hazard function, is random.

References

Aitkin, M., Laird, N., and Francis, B. (1983). An analysis of the Stanford Heart Transplant Data. *Journal of the American Statistical Association*, **78**, 264-292.

Akaike, H. (1974). A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, **19**, 716-723.

Anderson, P. K. (1982). Testing Goodness of Fit of Cox's Regression Model. *Biometrics*, **38**, 67-77.

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics*, **30**, 89-99.

Buckley, J., and James, I. (1979). Linear Regression with censored data. *Biometrika*, **66**, 429-436.

Chen, Y. Q. (2001). Accelerated Hazards Regression Model and its Adequacy for Censored Survival Data. *Biometrics*, **57**, 853-860.

Cook, R. J., and Lawless, J. F. (2002). The Analysis of Repeated Events. *Statistical Methods in Medical Research*, **11**, 141-161.

Cox, D.R. (1972). Regression Models and Life Tables. *Journal of the Royal Statistical Society*, **34**, 187-220.

Cox, D.R., and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.

Efron, B. (1977). Efficiency of Cox's likelihood function for censored data. Journal of the *American Statistical Association*, **72**, 557-565.

Fan, J., and Lin, F. (2002). Variable selection for Cox's proportional hazards model and frailty model. *The Annals of Statistics*, **30**, 74-99.

Fleming, T. R., and Harrington, D. P. (1999). *Counting Processes and survival analysis*. Wiley, New York.

Gehan, E. A. (1969). Estimating Survival Function from Life Table. *Journal of Chronic Desease*, **21**, 629-644.

Gray, G.R. (1992). Flexible methods that analyze survival data using splines. *Journal of the American Statistics Association*, **87**, 942-951.

Johansen, S. (1983). An extension of Cox's regression model. *International Statistics Review*, **51**, 165-174.

Hosmer, D. W., Lemeshow, S. (1989). *Applied Logistic Regression*. Wiley, New York.

Kalbfleisch, J. D. (1974). Some efficiency calculations for survival distributions. *Biometrika*, **61**, 31-38.

Kalbfleisch, J. D., and Prentice, R. L. (1973). Marginal Likelihoods Based on Cox's Regression and Life Table Model. *Biometrika*, **60**, 267-278.

Kalbfleisch, J. D., and Prentice, R. L. (2003). *Statistical Analysis of Failure Time Data*. Wiley, New York.

Kaplan, E. L., and Meier, P. (1958). Nonparametric Estimation from Incomplete Obsevations. *Journal of the American Statistical Association*, **53**, 457-481.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley, New York

Lee, E. T. (1980). *Statistical Methods for Survival Data Analysis*. Lifetime Learning Publications, Belmont, California.

Lee, E. T. and Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*. Wiley, New York.

Mantel, N. (1977). Test and Limits for the Common Odds Ratio of two-way Contingency Tables. *Journal of Statistical Planning Information,* **1**, 179-189.

Mantel, N. and Haenszel, W. (1959). Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. *Journal of the National Cancer Institute*, **22**, 719-748.

Miller, R. (1981). *Survival Analysis*. Wiley, New York.

Myers, M., Hankey, B.F., Mantel, N. (1973). A Logistic-Exponential Model for Use with the response-time data involving regressor variables. **Biometrics**, **29**, 257-269.

Peto, R., and Peto, J. (1972). Asymptotically Efficient Rank Invariant Procedures. *Journal of the Royal Statistical Society*, **135**, 185-207.

Prentice, R. L. (1973). Exponential Sruvivals with Censor and explanatory variables. *Biometrika*, **60**, 279-288.

Prentice, R. L. (1976). Use of Logistic regression models in Retrospective studies. *Biometrics*, **32**, 599-606.

Prentice, R. L., and Kalbfleisch, J. D. (1979). Hazard rate models with covariates. *Biometrics*, **35**, 25-39.

Rao, C. R. (1973). *Linear Statistical Inference and its Application*. Wiley, New York.

Schwarz, G. (1978), Estimating the Dimension of Models. *The Annals of Statistics*, **6**, 461-464.

Segal, M.R., and Bloch, D. A. (1989). A comparison of estimated proportional hazard models and regression trees. *Statistics in Medicine*, **8**, 539-550.

Sparling, Y. H.,Younes, N. and Lachim, J. M. (2006) Parametric Survival Models for interval-censored data with time-dependent covariates. *Biostatistics*, **4**, 599-614.

Tarone, R. E., and Ware, J. (1977). On distribution-free tests for equality of survival distribution, *Biometrics*, **64**, 156-160.

Yu, S. Y., and Voit, E. O. (1995). A simple flexible failure model. *Biometrical Journal*, **37**, 595-609.

Zhao, Q., and Sun, J. (2004). Generalized log-rank test for interval-censored failure time data. *Statistics in Medicine*, **23**, 1621-1629.