First Term Probation: Models for Identifying High Risk Students[1]

Alejandro Scalise, Mary Besterfield-Sacre, Larry Shuman and Harvey Wolfe
School of Engineering- University of Pittsburgh

Abstract - *EC 2000 has heightened the awareness among engineering faculty of the importance of student retention, especially the retention of first-year students. Previous research has found that students placed on academic probation after their first term have a high probability of leaving engineering prior to graduation. Using eight years of data, we examine the influence of the student's initial preparedness, attitude toward his/her chosen career, and self-assessed confidence in factors such as confidence in study habits and communication skills on first term probation and retention. Logistic regression approaches are used to develop models to identify those engineering students most likely to be placed on first-term probation. These models have enabled us to determine the factors that most influence first term probation and to better identify students who require early interventions if they are going to successfully complete the engineering curriculum. We propose that such models are valuable tools for engineering educators concerned about retention and graduation rates.*

Introduction

The Need to Analyze Retention

Student retention is one of the major challenges facing currently facing engineering faculty and administrators [1]. Almost 50% of the students entering an engineering program leave before graduation with a large part of this attrition occurs during the first year [2][3]. With the introduction of ABET's EC-2000 criteria that requires engineering programs to assess and improve student continuity and stability, retention's importance has increased even more [4].

There have been a number of efforts to reduce attrition [5][6]. Some were directed at analyzing the response to curricular changes while others were oriented to the development of models to predict attrition [7][8]. These applications are relevant because factors that affect retention can be identified through the model building process and the resultant models can be used to identify students most likely to leave prematurely. This allows interventions to be designed and targeted to assist those students that need the most help.

A substantial percentage of the students who leave engineering before graduation had been placed on first term probation [3]. Reasons for analyzing engineering student retention include:

- Increasing demand for engineers by the industry.
- Decreasing enrollment in engineering programs.
- Large attrition rates in engineering programs.
- Recruiting costs are greater than retention costs.
- Retention rates may be used for benchmarking [9].

A purpose of ABET's site visits is to assess difficult to evaluate qualitative factors including " the stability and continuity of the faculty and the students" [4].

Retention studies have analyzed the need for specific courses [10][11]; subjective factors such as initial preparedness, attitudes and confidence measures, gender and ethnicity; and the effect of program factors provided such as quality of teaching and advising, available resources, and financial aid provided. The causes of attrition include [12][13]:

- Losing interest in engineering and finding more interest in other majors.
- Poor teaching by engineering faculty.
- Overwhelming pace and load of engineering programs.
- Grading systems in engineering are discouraging.

Our objective is to better understand the factors that affect first term probation, so that better interventions can be designed and implemented. We are interested in determining how retention is affected by subjective factors that can be measured before freshmen begin their courseware and encounter new experiences and influences. By improving the first term performance of students at risk, the chance of successful completion of the engineering program are increased.

We present Logistic Regression models developed at the University of Pittsburgh School of Engineering to predict first term probation. Information about the student's prior academic achievement and abilities as given by high school rank and SAT scores and his/her attitudes and confidence as measured by the Pittsburgh Freshman Engineering Attitudes Survey© are used to develop these models. All required data are collected at during registration enabling predictions to be done before the student begins his/her first freshman classes.

Models for Retention

Aitken [14] highlights the importance of knowing the factors that affect student retention, and the need for a "comprehensive model to capture the structure of the process" and to effectively predict retention. He developed a model of retention as a function of the student's academic performance, academic satisfaction, and living satisfaction. In turn, each one of these endogenous variables is a function of other exogenous variables including SAT scores, high school rank, GPA, perceived GPA (expected), quality of academic advising, quality of instruction, student's personality, motivation. Many of these variables were obtained from a ques-

[1] This paper supported in part by National Science Foundation grant: EEC-9872498, *Engineering Education: Assessment Methodologies and Curricula Innovations* and Engineering Information Foundation grant 98-4.

tionnaire that the students completed at the end of the freshman year. Knowing these factors allows measures to be taken in the future to improve retention. One drawback of this model is that some of the required information can only be obtained towards the end of the first year. Thus, interventions are not likely to be applied to the most high risk group before many have left engineering.

Jensen [15] used path analysis to model the effect of financial aid on persistence. The number of semesters a student attended school was used as dependent variable; independent variables included measures of socioeconomic status, academic background, and financial aid received. The results indicated that financial aid does effect persistence, and that the refusal of financial aid has a negative influence. Again, some of the required information may not be available until after the academic year has begun.

Logistic Regression models were used in one study at the University of Iowa [9] to predict the risk level of attrition. Four possible categories were analyzed: low, mild, medium and high risk. Factors included were: ACT scores, high school rank, number of semesters of physics and social science in high school, marital status, and race. Student's attitudes were not considered.

The Need to Analyze First Term Probation

Being placed on first term probation influences the student's future success – a large number of these students leave the program, even after moving off of probation. Table 1 shows that the percentage of our students leaving the School of Engineering is much higher for those students placed on first term probation. That is, for freshmen who entered in the fall of 1995, 59% of those placed on first term probation subsequently left engineering during the succeeding four years compared to 24% of those who did not go on probation during their first term.

YEAR	First Term Probation	
	No	Yes
1995 - 1996	24%	59%
1996 - 1997	14%	54%
1997 - 1998	13%	33%
1998 - 1999	5%	12%

Table 1- Students Transferring Out of Engineering

Table 2 shows that the percentage of students leaving the program before the end of the second term (freshman year) is much larger among students placed on first term probation. The last two columns show that their academic performance continues to be worse than that of students not placed on probation.

A longitudinal study at Purdue [3] confirmed the previous observations and showed the high correlation between first term GPA and retention. A conclusion was that first term probation is a more accurate predictor of retention than SAT scores. Also demonstrated was that a good performance in basic mathematics and physics courses gave students the needed confidence to complete their program. These observations can be translated directly into freshman year interventions to prevent high risk students to go on probation.

Table 2: Comparative Performance of Freshmen

YEAR	First term probation	Left before the end of the second term	Second Term probation	
			No	Yes
1995 - 1996	No	8.2%	76.8%	14.9%
	Yes	29.1%	23.6%	47.3%
1996 - 1997	No	5.6%	82.3%	12.1%
	Yes	26.9%	31.3%	41.8%
1997 - 1998	No	.3%	89.1%	10.5%
	Yes	9.1%	49.1%	41.8%
1998 - 1999	No	6.3%	79.5%	14.2%
	Yes	29.7%	26.4%	44.0%

Confidence and Attitude Measures

Students' performance and retention are related to the opinion students have of themselves [8]. The Pittsburgh Freshman Engineering Attitudes Survey© (PFEAS) is used to measure freshmen attitudes toward an engineering career and the student's self assessed confidence in such areas as study habits and communications skills. Thirteen attitude and confidence measures have been derived from this instrument's fifty items using statistical methods [2]. Besterfield-Sacre, et al. used these measures in models to predict freshman attrition, thus allowing for the implementation of interventions that would reduce the flow of good students from engineering [8]. The models predicted the likelihood of a student of leaving engineering in good academic standing. Interventions included curricular changes and special classes to strengthen confidence and change attitudes.

Logistic Regression

Logistic Regression (LR) is a special case of regression, with a binary outcome – one or zero – that indicates the occurrence or non-occurrence of a given event or result, or the presence or absence of a expected characteristic. The coefficients for such a model can be used to analyze the modification in the outcome's odds for a change in one of the predictor variables. Thus, with a logistic regression model the investigator can study the likelihood of the occurrence of an

[2] These measures are: 1) general impressions of engineering, 2) financial influences for studying engineering, 3) perception of how engineers contribute to society, 4) perception of the wok engineers do and the engineering profession, 5) enjoyment of math and science courses, 6) engineering perceived as being an "exact" science, 7) family influences to studying engineering, 8) confidence in basic engineering knowledge and skills, 9) confidence in communication and computer skills, 10) adequate study habits, 11) working in groups, 12) problem solving abilities, 13) engineering ability.

event given a set of values for the predictor variables and how each predictor variable affects such likelihood.

An operation called "logit" is applied to the event's probability, and is represented as a linear function of the predictors. The logit is regressed against the predictors. The formula for this function of the event's probability is given by: $\text{logit}(p) = \ln(\frac{p}{1-p})$, where p is the probability of the event.. When the inverse operation is applied to the logit, an S-shaped curve is obtained which represents the probability of the event occurring [16]. In contrast to linear regression, the coefficients for the LR model have to be solved iteratively and the likelihood evaluated at each iteration.

Methodology

Data

Data from freshman classes for 1995-96, 1996-97, 1997-98, 1998-99 and 1999-00 were used in this study. Subjective and objective measures were collected for each student including initial preparedness, ability, attitude and self assessed confidence. The relationship of these factors to first term probation was then analyzed. Variables included:

- SAT score was used to measure initial preparedness.
- High school rank was used to represent students' relative academic achievement.
- The thirteen PFEAS factors were used to measure attitude and confidence at entry.
- Gender was included to see if it were a factor.
- A zero-one dummy variable (Impact) indicated if a student participated (1) in the summer bridge program (which was primarily for minority students).
- First Term QPA (the student's QPA at the end of the fall term) measured the degree of achievement.
- A dummy variable (Observed Probation or OP) was created to distinguish students placed on probation (1) from those who were not (0). Students with a QPA below 2.00 received a value of one.
- A dummy variable (Predicted Probation or PP) distinguished between students predicted to be placed on probation (1) from those who were not (0).

The data were split into four sets. The first was used for model construction. It included 66 % randomly chosen cases from 1995-96 through 1997-98. The remaining 33% of the cases became the test set, and was used to fine-tune the model and test its performance. The data from 1998-99 were used as an independent test set and data from class 1999-00 were used to predict how the new students would perform before they completed their first term.

Nomenclature

The following nomenclature is used:
"**Positive**": Students with First Term QPA below 2.00 (OP = 1); first term probation.
"**Negative**": First Term QPA of 2.00 or better (OP = 0); good academic standing.
"**True positive**": Students who are predicted to and did go onto first term probation (PP = 1, OP = 1).
"**False positive**": Students who are predicted to but did not go onto first term probation (PP = 1, OP = 0).
"**True negative**": Students who are predicted to but did not go onto first term probation (PP = 0, OP = 0).
"**False negative**": Students who are predicted not to but did go onto first term probation (PP = 0, OP = 1).

Statistical Analysis

Non-parametric tests were used to analyze the relationship between our predictor variables and the outcome (First Term Probation). Parametric methods were discarded because most of the attitude and confidence measures do not follow the normal distribution. Hence using these methods would introduce a serious bias in the results.

A Kruskall-Wallis non-parametric test was used to identify factors whose values differentiated between the students placed on first term probation (positives) and those who were not (negatives). The p-value indicated if there was a significant difference in between the two groups. An univariate analysis complemented this analysis; here Logistic Regression models (one for each candidate independent variable) were developed with one predictor variable regressed against the outcome. Once a model was obtained, the significance of the predictor's coefficient indicated the strength of the relationship between the variables. Only those candidates with a small p-value were considered in the first steps of the model development.

Model Development

The resultant variables from the first analysis were used to develop a first Logistic Regression model[3]. After obtaining an initial model, the relationship between the predictors included in the model and the logit function is analyzed in order to identify any non-linearity. When a relationship is not linear, the model can be improved by applying transformations to the predictor variables, e.g., square root, log, or by creating categorical variables where each category represents an interval of the original predictor.

The next step in the modeling process is to find any important, but presently not included predictors. Variables that were not selected in the initial analysis may become important predictors when other variables are either added or transformed. Thus, a second stepwise process is performed using variables that the analyst considers as still potentially important factors. After this step, the predictor variables

[3] The forward stepwise method, an iterative procedure that fits the model by evaluating variables to be included and removed at each step, was used. At each iteration, all candidate predictors not yet in the model are tested one at a time, and the improvement in the likelihood function obtained. The one that produces the best model is kept. Next each variable already in the model is removed, and the change analyzed to identify the best fit at this iteration. The process is repeated until no more variables can be added or removed.

should be analyzed for possible interactions among them, again using a new stepwise process.

An improved model should results which captures some of the factors that influence the outcome since the model will account for only a portion of the outcome's variability. Unexplained variability may be due to missing predictors (information that has not been collected because of the difficulty of collection or because it has not been identified) and the inherent randomness of the outcome (given two students with the same initial conditions, one may be placed on first term probation while the other may not).

Finally, diagnostics should be performed to identify data points that are poorly fit, may be considered outliers, or may have an excessive influence in the determination of the model's coefficients. (Elements such as deviance, leverage and DfBetas are useful in performing these diagnostics.)

Model Evaluation

The resultant model should be evaluated using several types of analysis. The goodness of fit is analyzed to determine how well it represents the data. We used the Hosmer and Lemeshow Goodness of Fit Test, a version of a chi-square test where the p-value indicates the significance of the difference between prediction and actual. Another measure predictive ability is the percentage of probation students identified. This is done by using the model with the test set (i.e., data that was not used for model development) to obtain the predicted probability of first term probation. The predicted probation group is compared to the actual one. Each prediction is classified into one of the four groups: true positive, false positive, true negative and false negative defined above. The percentage of true positives with respect to the total number of students on probation is used to evaluate the model's ability to identifies students at risk.

Table 4. Statistical Analysis

Variable	Mann-Whitney Test	Univariate Analysis
SAT Total	*	*
HS Rank	*	*
Gender	*	
Impact	*	
Career		
Jobs		
Society		
Perception		
Math		
Exact Science		
Family Influence		
Basic Engineering	*	**
Communications		
Study	*	*
Groups		
Ability		
Engineering	*	

* Significant at the 0.05 level
** Significant at the 0.10 level

Also, the ratio between the number of false positives to the number of true positives gives a measure of the prediction error. This ratio is called the model "cost." A value of 2.0 for this ratio indicates that we incorrectly predict two students to go on probation for each one correctly predicted. Because of the unknown factors that may influence first term probation, we consider a cost below 2.0 to be acceptable.

Model Tuning

The first test set was used to fine-tune the model. While, the model's output is the probability of first term probation, we also use it to classify students into two groups: those most likely to be placed in such condition and those that are not likely. To do this, we choose a threshold or limit probability such that a student whose probability is larger than or equal to the threshold is "predicted" to go onto probation. To best choose this limit, a test set was used to evaluate different thresholds relative to classification rate and the "cost."

Results

Statistical Analysis

The results of the non-parametric tests are summarized in Table 4 with the results of the univariate analysis. These analyses indicate significant differences between positives and negatives in SAT, Rank, Gender, Impact, Confidence in Basic Engineering Knowledge and Skills, Confidence in Study Habits, and Confidence in Engineering Ability.

First Complete Model

Two models were developed. The first one, "Model 1," includes SAT, square root of high school rank and a categorical variable, which measures students' self-assessed confidence in their current study habits. The model is given by:

$$\begin{aligned}\mathrm{logit}(p) = \beta_0 &+ \\ &\beta_1 * \mathrm{SAT} + \\ &\beta_2 * \sqrt{\mathrm{Rank}} + \\ &\beta_{3i} * \mathrm{Adequate\ Study\ Habits}_i\end{aligned}$$

This linear formula is used to calculate the logit function where β_0 is the interception, and the other betas are the coefficients for each of the independent variables. For SAT and square root of rank, the betas represent the increase in the outcome for a unit increase in the variable. For "Adequate Study Habits," β_{3i} represents the increase in the logit when the student's confidence falls into the i^{th} category (I = 1, 2, or 3). The baseline for this model is given by SAT = 0, rank = 0, and Study = 4.

The probability of first term probation is derived by: $p = 1/(1+e^{\mathrm{logit(p)}})$. Given this probability, a threshold limit is applied to predict if the student is a candidate for first term probation. We have chosen a limit of 0.15; a student with probability larger than or equal to this value is predicted to be positive. Otherwise, he /she is predicted to be negative.

For some students the SAT score and/or the high school rank were not available. To resolve this problem, a second model ("Model 2") was developed which did not include these two variables. The logit for this model is:

$$
\begin{aligned}
\text{logit}(p) = \beta_0 + {} & \\
& \beta_1 * \text{Impact Program} + \\
& \beta_2 * \text{Engineering Abilities} + \\
& \beta_{3i} * \text{Adequate Study Habits}_i + \\
& \beta_4 * \text{Problem Solving Abilities} + \\
& \beta_5 * \text{Impact Program} * \text{Engineering Ability}
\end{aligned}
$$

As in the previous case, probation can be predicted using the calculated probability and a threshold limit of 0.15.

Table 5 – Coefficients and Odds Ratios

	Variable	Coefficient	Odds Ratio
Model 1	**Sat**	-0.01	0.99
	Sqrt(Rank)	3.4	29.9
	Study 1	1.7	5.48
	Study 2	1.32	3.76
	Study 3	0.61	1.84
	Constant	4.37	
Model 2	**Study 1**	2.77	16
	Study 2	2.31	10.1
	Study 3	1.5	4.49
	Impact 1	6.32	556
	Engineering	0.88	2.4
	Ability	-0.78	0.46
	Interaction	-1.57	0.21
	Constant	-3.96	

A combination of both models is used in order to evaluate the largest number of students. Model 1 is used if all required information is available. Model 2 is used only where Model 1 is not applicable due to incomplete information. Table 5 summarizes the values of the betas for both models, together with the odd-ratios, discussed below.

Model Evaluation

Hosmer and Lemeshow Goodness of Fit Test was applied to both models resulting in p-values of 0.5543 for the first and 0.7622 for the second. This implies that there is no significant difference between the positives and negatives predicted by the models and the actual results.

Table 6 summarizes the classifications results obtained from applying both models to the 1999-00 entering freshmen. It shows that 54 of 63 positives (first term probation) are identified (86%). In contrast, another 94 were incorrectly predicted to go on probation. That is, there is a cost of 1.74 false predictions for each correct prediction. The percentage of true positives could be increased to 100% by decreasing the threshold. However, the incremental cost in terms of increased false predictions would be unacceptable. Table 7 shows that the model performs equally well for special student populations; women, minorities and those in our pilot integrated curriculum, indicating the models' robustness.

Table 6: Classification of Results

	Predicted Probation	
Observed Probation	**Negative**	**Positive**
Negative	174	94
Positive	9	54

Table 7: Predictions for Special Populations

Special Case	True Positives	Cost
Male	86%	1.75
Female	83%	1.7
Integrated Curriculum	100%	1.86
Minority Program	85%	1

A concern in applying these models is the large number of students who are incorrectly predicted to go on probation (false positives). This is partially due to the low threshold chosen (0.15). However, these students do enter with some academic risk. While they may not go on probation, our historical data indicates that their aggregate performance will be lower than the performance of the true negative group. Figure 1 shows average first-term QPA for the four predictions categories. Note that false positives had a smaller average QPA than did true negatives. The conclusion is that that students predicted to be positive will not perform as well, in general, as those students in the true-negative group, even if they do not go on probation.

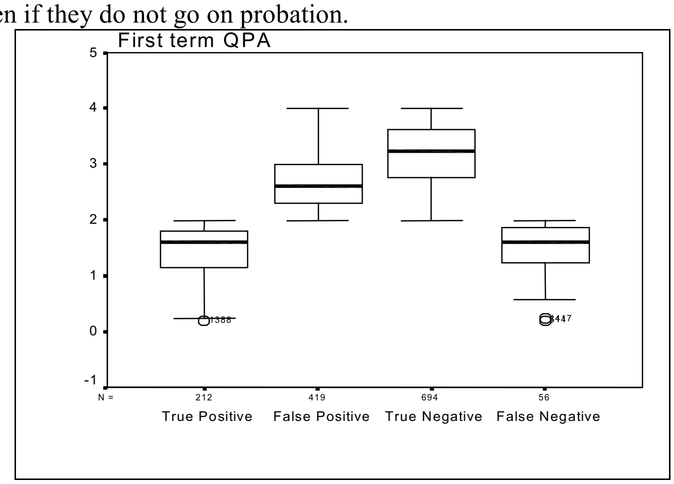

Figure 1-First Term GPA Distribution by Classification

Conclusions

The conclusions that can be drawn from this study are:
First, there is a clear need to understand first term probation and the factors that affect it, and to identify risk students. Second, factors that may affect the risk of a student being placed on first term probation can be identified and weighted by "importance" in relation to the probability of being placed into probation. It is important to identify and classify these factors in order to design more effective interventions; e.g., those that address problems such as the lack of adequate

study habits, poor initial preparation in math, and the need to reinforce abilities required to succeed in engineering such as problem solving skills.

The most important predictor factors that we identified (Model 1) included:

- Preparation, as measured by SAT.
- Academic ability as measured by high school rank.
- Confidence in study habits, as measured by the PFEAS.

Other factors (Model 2) include:

- Confidence in problem solving abilities (PFEAS).
- Confidence in engineering abilities (PFEAS).
- High risk admission through a bridge program.

A third class of factors was identified through the univariate analysis but not included in the models. The only variable in this class – low confidence in basic engineering knowledge and skills – while significant by itself was not significant when used with the predictors included in the models.

Finally, there is a class of potential predictors that were found not to affect the probability of being placed on first term probation. These included: gender; general impressions of engineering; financial influences for studying engineering; perception of how engineers contribute to society; perception of the work engineers do and the engineering profession; enjoyment of math and science courses; engineering perceived as being an "exact" science; family influences to studying engineering; confidence in communication and computer skills; working in groups.

The probability of first term probation is effected by predictors' values. The betas represent the logarithm of the odds ratio between two different states of a variable (see Table 5); e.g., a unit change in SAT represents an odds ratio of 0.99, what implies that odds decrease 1% for each unit increase in the SAT score. Also, a unit change in rank (the full range for this variable) increases the odds by almost 30 times. It is interesting to observe that a 100 units decrease in SAT or a 0.05 unit increase in square root of rank (0.0025 increase in rank) give an odds-ratio of 2; i.e., the odds doubles when one of these undergoes a relatively small change.

An advantage of these models is that they allow us to direct our interventions only on those freshmen who are most likely to go on first term probation. That is, the models allow us to focus on slightly over a third of the freshman class, and in that manner capture 86% of those who went on first term probation. This ability to hone in on those most in need reduces the costs of applying the interventions. On the other hand, a possible disadvantage of a directed intervention is that the recipients may feel that they are being discriminated when placed in special classes or workshops. We are currently wrestling with these issues as we plan the implementation of these models for the 2000-01 entering freshmen.

References

1. Besterfield-Sacre, M, Shuman, LJ, Atman, CJ, "Perception versus Performance: The Effects of Gender and Ethnicity Across Engineering Schools," Frontiers in Education, November 1998.
2. Felder, R, Felder, G, Dietz, E, "A Longitudinal Study of Engineering Student Performance and Retention. V. Comparisons with Traditionally-Taught Students," Journal of Engineering Education, Vol. 87 No. 4, October 1998, pp. 469-480.
3. Budny, D., LeBold, W., Bjedov, G. "Assessment of the Impact of Freshman Engineering Courses," *Journal of Engineering Education*, Vol. 87 No. 4, October 1998, pp. 405-411.
4. Engineering Accreditation Commission,_Accreditation Policy and Procedure Manual, (Baltimore, Maryland: Accreditation Board for Engineering and Technology, Inc. November, 1998).
5. Jensen, E., "Student Financial Aid and Persistence in College," *The Journal of Higher Education*, Vol. 52, No. 3, May-June 1981, pp. 280-294.
6. Brainard, S., Carlin, L., "A Six Year Longitudinal Study of Undergraduate Women in Engineering and Science," *Journal of Engineering Education*, Vol. 87 No. 4, October 1998, pp. 369-375.
7. Aitken, N., "College Student Performance, Satisfaction and Retention: Specification and Estimation of a Structural Model," *The Journal of Higher Education*, Vol. 53, No. 1, 1982, pp. 280-294.
8. Besterfield-Sacre, M., Atman, CJ, and Shuman, LJ. "Characteristics of Freshman Engineering Students: Models for Determining Student Attrition and Success in Engineering," *Journal of Engineering Education*, Vol. 86 No. 2, April 1997, pp. 139-149.
9. Moller-Wong, C., Eide, A., "An Engineering Student Retention Study," *Journal of Engineering Education*, Vol. 86 No. 4, January 1997, pp. 7-15.
10. Hatton, Deborah, Wankat, Phillips, Lebold, William, "The Effect of an Orientation Course on the Attitudes of Freshmen Engineering Students," Journal of Engineering Education, Vol. 87 No. 1, January 1998, pp. 23-27.
11. Hoit, M and Ohland, M. "The Impact of a Discipline-Based Introduction to Engineering Course on Improving Retention," *Journal of Engineering Education*, Vol. 87 No. 1, January 1998, pp. 79-85.
12. Seymour, E and Hewitt, N. *Talking About Leaving: Why Undergraduates Leave the Sciences*. Westview, 1997.
13. Tinto, V. *Leaving College: Rethinking the Causes and Cures of Student Attrition*, second edition, University of Chicago Press, 1993
14. Aitken, N., "College Student Performance, Satisfaction, and Retention," *Journal of Higher Education*, Vol. 53 No. 1, 1982, pp. 32-50.
15. Jensen, E., "Student Financial Aid and Persistence in College," *Journal of Higher Education*, Vol. 52 No. 3, 1982, pp. 280-294.
16. Hosmer, D., Lemeshow, S., *Applied Logistic Regression*. John Wiley & Sons, New York, 1989.