Modeling consumer acceptance probabilities

L.C. Thomas *, Ki Mun Jung [b], Steve D. Thomas [a], Y. Wu [a]

[a] *School of Management, University of Southampton, Southampton SO17 1BJ, UK*
[b] *Department of Informational Statistics, Kyungsung University, Busan 608-736, South Korea*

Abstract

This paper investigates how to estimate the likelihood of a customer accepting a loan offer as a function of the offer parameters and how to choose the optimal set of parameters for the offer to the applicant in real time. There is no publicly available data set on whether customers accept the offer of a financial product, whose features are changing from offer to offer. Thus, we develop our own data set using a fantasy student current account. In this paper, we suggest three approaches to determine the probability that an applicant with characteristics will accept offer characteristics using the fantasy student current account data. Firstly, a logistic regression model is applied to obtain the acceptance probability. Secondly, linear programming is adapted to obtain the acceptance probability model in the case where there is a dominant offer characteristic, whose attractiveness increases (or decreases) monotonically as the characteristic's value increases. Finally, an accelerated life model is applied to obtain the probability of acceptance in the case where there is a dominant offer characteristic.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Student bank account; Acceptance probability; Coarse classifying; Logistic regression model; Linear programming; Accelerated life model

1. Introduction

Forecasting financial risk in consumer lending has over the last thirty years become a major growth areas (Rosenberg & Gleit, 1999; Thomas, 2000; Thomas, Edelman, & Crook, 2002). The main approaches are credit scoring and behavioural scoring which are based on statistical or operational research methods of classification. The statistical methods include discriminant analysis, logistic regression, classification trees and survival analysis, while the operational research techniques include linear programming. Discriminant analysis (which is equivalent to linear regression) was proposed by Fisher (1936) as a discrimination and classification tool and was one of the first methods applied to building credit scoring models by discriminating between those loans which in the past had defaulted and those which had not defaulted. Logistic regression is a related statistical modeling method which is now widely used and Wiginton (1980) was one of the first to describe the results of using it in credit scoring. The application of survival analysis for building credit scoring models was introduced by Narain (2004) and developed further by Stepanova & Thomas (2002). Mangasarian (1965) was the first to recognize that linear programming could also be used in classification problems

where there are two groups (defaulters and non-defaulters in this case), and this work was extended by Freed & Glover (1981); Hand (1981). In all cases, the approach is to take a sample of previous applicants and to find the best way of separating those who default from those who do not default.

In the same way as assessing default risk, statistical and operational research method can be used to determine the probability that an applicant with certain characteristics will accept the offer of a loan (or respond to some direct marketing literature). Here, the two groups are those who take up the offer and those who do not or in the case of direct marketing those who respond to the mailing and those who do not. Response scorecards are widely used in marketing but the acceptance scorecards have had less visibility. However, with the lenders slowly changing their objectives from minimising the risk of default to maximizing the profitability of the loans they offer, estimating this probability of acceptance is becoming more important.

Currently, there are two significant changes in the consumer lending process which have added to the need to estimate probability of acceptance but for a more complex problem. The first is that instead of a lender having a specific product to offer— a credit card with a given interest rate for example,— increasingly they can offer a generic product but where the features of the product can vary from consumer to consumer. It has always been the case that the overdraft limit varied but with the advent of risk based pricing, the interest rate offered is also beginning to be varied from consumer to consumer. Similarly in

---

* Corresponding author.

credit cards, other features that can be varied would be an annual fee, an initial discount offer, a points scheme and free travel insurance. Other financial products also have variable features. For bank accounts, these would be the overdraft limit, an ATM card, interest paid when in credit, no fee foreign exchange while for mortgages one can vary the loan amount, the interest rate, the connection if any to the official interest rate, a initial discount on the rate, and whether there is cashback. Customer relationship management implies that the lender wants to tailor these features to the customers' requirements and so make it more likely that they will accept and stay with the product.

The second change is that the newer communication and marketing channels, which can be used for applying for loans,—the internet and the telephone for example—allow for the features of the offer to be adjusted during application process. A lender offers the generic product, the credit card say, but during the application process when obtaining the applicants' information in order to check their default risk, the lender has the opportunity to tailor the features of the product so as to make the applicant more likely to accept the product, if they are offered it. This contrasts with the traditional form filling approach where there is really no interaction in the application process between potential borrower and lender.

This means it is now important for a lender to be able to assess the likelihoods of a particular consumer accepting each of the variants of the particular product that he is offering. Moreover, this has to be done in real time during the application process so that the most 'suitable', product can be offered at the end of the process. One may ask why does the lender not show all the variants of the product to the consumer and let him choose which he wants. In many retail environments, e.g. clothes, furniture, this is a sensible strategy but in the financial sector there are three reasons why this is not the way things are developing. Firstly, customer are likely to chose the most 'unprofitable' product for the lender. Secondly, the prestige of a financial organisaton is lowered if it implies it does not know what is the best product for its customer. Thirdly, with so many combinations, the customers will be spoiled for choice, which unsettles some consumers especially in the financial area.

So in this paper, we investigate how to estimate the likelihood of a customer accepting a loan offer as a function of the offer features. Then the lender can decide which acceptable combination of features is optimal to offer the applicant, and both these calculations must be done during the application process. Here, an acceptable combination may be one that is profitable to the lender and the optimality is judged in terms of maximizing the probability of the offer being accepted. Alternatively, an acceptable combination might be one that meets the customer's minimum requirement and the optimality is to maximize the lender's profit.

There has been a significant amount of work recently on how management science approaches can support marketing through the internet, for example the special issues edited by Geoffrion & Krishnan (2003); Kannan and Rao (2001). Montgomery (2001) took as one of his examples of real internet applications the use of conjoint analysis and logistic regression to identify the importance of the separate components of a product in their utility to the customer. He applied it to identifying the importance of the item price, the shipping price, the sales tax and the delivery time for an internet bookseller. This has some connection with the first approach in this paper but whereas our objective is to decide in real time on a suitable price for the product, his work was more on identifying the importance of the features. A second example in that review was the idea of pricing using versioning in which variants of a product are sold at different prices. Rossi, McCulloch and Allenby (1996) sought to estimate the price sensitivity of a customer to the different alternative versions of the item using that customer's previous purchase history and used that to target the distribution of a coupon which discounted the price of the item.

The internet is also leading to personalization as a way of using data on customers to give them a better choice. Personalization is the process of using this information to develop a targeted solution to the customer, which we are suggesting is the way in which financial institutions will move. Murthi & Sarkar (2003) survey the work in this area looking at the strategic implications of personalization and the standard techniques that can be used to model the information. These include the classification techniques which were previously developed in the credit scoring area and consumer choice models. They make the point that in the taxonomy of price discrimination the ability to charge different prices for the same product to different customers is first degree price discrimination while offering different products at different prices is second degree price discrimination. They point out that in traditional markets first degree discrimination is not practical but in the internet age it becomes a possibility What we are proposing here is a tailoring both of product and price which is neither first nor second degree discrimination. Karuga, Khraban, Nair, and Rice (2001) develop a genetic algorithm approach to deciding which advertisements to target to a customer using the internet and how to schedule the advertisements. They use a linear programme to identify the 'effectiveness' of the features that make up the advertisement but it differs from the way we use linear programming and their objective is to identify which advertisements and which sequence of advertisements produces the highest response rate, not how to adjust price and product features so as to maximize the probability of acceptance. Raghu, Kannan, Rao, and Whinstom (2001) consider the next step in the customization problem—namely how to dynamically update the model. They consider the use of questionnaires to dynamically update the consumers' preferences.

Since, the use of these techniques is so new and so little has been done on jointly adjusting price and product features, one problem is the data. There is no publicly available data set on whether customers accept the offer of a financial product - the features of which are changing from offer to offer. Thus, we have developed our own data set using a fantasy student current account (FSCA) which is a web-based application form targeted at students applying for a bank account. Although it is not argued that this data reflects what will happen if students are applying

for a real bank account, it does enable us to investigate different approaches to building acceptance functions.

This paper suggests three approaches to determine the probability that an applicant with characteristics x will accept an offer with characteristics o using the FSCA data. Section 2 describes the FSCA data, which is used to build the model of consumer acceptance probabilities. In Section 3, a logistic regression (LR) model is used to obtain an acceptance probability estimate. In Section 4, a linear programming approach is used to build an acceptance probability model. To build such a model, we assume there is a dominant offer characteristic, where the attractiveness of the offer increases (or decreases) monotonically as this characteristic's value increases. The idea is that given the other offer and applicant characteristics one can identify the value of this dominant characteristic at which this applicant would accept the offer. Section 5 extends this idea by saying this cut-off value between accepting and rejecting of the dominant characteristic need not be fixed but can be described by a probability distribution. One can then use an accelerated life (AL) approach to obtain this distribution. This probability distribution of acceptance reflects the uncertainty consumers have in choosing between closely matched offers as well as the changes in the individual's mood, the environment where the offer is made and the knowledge of competing offers which occur between identical offers being made to customers with identical applicant characteristics. In applying the accelerated life model, one has double censoring, since if a customer accepts an offer with a particular overdraft limit one only knows the minimum acceptance value is below this value, while if he rejects an offer, one only knows the minimum acceptance value is above this offer value.

2. Fantasy student current account data

In order to model the likelihood of consumers accepting the offer of a financial product-the features of which are changing from offer to offer, we developed our own data set using a fantasy student current account (FSCA). The FSCA consists of a website which closely follows the application forms for student bank accounts which are used by the major UK banks. The website consists of three pages. The first is an application form for a FSCA, which is similar to the bank account most students in the UK use for their money transactions and their borrowing. The questions are created by looking at the application forms of ten UK lenders including the four major UK retail banks. The questions asked in the form are described by the application characteristics in Table 1, where the first eight concerned demographic and financial information, and the remaining ten addressed interests (some banks did have one or two such marketing oriented questions in their application forms). The second web page makes the offer of an account and outlines the features that were part of the offer. There were six features of the account that could be changed from offer to offer—given by the offer characteristics of Table 1.

In order to obtain a reasonable spread of offer combinations each applicant was randomly put into one of four offer

Table 1
Application and offer characteristics used in the model

Characteristics	Description
Application characteristics	
Age	Age of applicant
Sex	Sex of applicant
Status	Marital status of applicant
Num_children	Number of children
Num_cards	Number of credit card
Wage	Some income from wage
Loan	Some income from loan
Contribution	Some income from parental contribution
Travel	Interest in travel, true/false
Music	Interest in music, true/false
Cars	Interest in cars, true/false
Cinema	Interest in cinema, true/false
Sports	Interest in sports, true/false
Clubbing	Interest in clubbing, true/false
Beer	Interest in beer, true/false
Country western	Interest in C and W music, true/false
DIY	Interest in DIY, true/false
Gardening	Interest in gardening, true/false
Offer characteristics	
Overdraft	Overdraft limit, five choices
Creditcard	Credit card included with account, four choices
TM	No fees on ordering foreign currency for travel, yes/no
Insurance	Discounts on insurance, four choices
Interest	Interest paid when account in surplus, four choices
Introductory	Introductory free gift, 10 choice

categories. In three of these categories, everyone in that category received the same fixed offer varying from £1250 to £1800 overdraft limit. In the fourth category which had the largest probability the offer was given by one of 42 nodes of a decision tree arrived at by splitting on the applicants' characteristics. The decision tree was constructed subjectively, using obvious associations and a desire to produce a wide spectrum of offers. When the offer was made on the second page, the applicant had to submit whether they accepted or rejected the offer. This is the outcome that the subsequent models seek to estimate. The final page asks applicants to rate the importance of the offer characteristics in their decision and also how they feel about a bank making different offers to different people.

This web page (www.management.soton.ac.uk/staff/fair-isaacs/win.asp) is available for all to use but it was publicized widely to the first year students at the University of Southampton in the UK, with regular prize winning draws for those who had completed the application process. Although there was no guarantee the acceptance/rejection decision was the one they would have made to a real account they had all recently opened such accounts and so were well aware of the product and the features offered by the various banks.

A data set of the application and offer characteristics for 331 applicants was obtained from the website. There are 18 applicant characteristics and six offer characteristics. In Sections 3–5, we will deal with three approaches to determine the probability that an applicant with applicant characteristics x will accept an offer with features o using this FSCA data. In each case, we build the model on 265 of the cases in the sample

and test it on the remaining 66 cases (so this holdout sample is 20% of the population).

3. Logistic regression based acceptance probability approach

Logistic regression is a widely used statistical modeling method in which the probability of a dichotomous outcome is estimated. In general, the logistic regression model has the form

$$\log\left[\frac{p}{1-p}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k = x\beta, \quad (1)$$

where p is the probability of the outcome of interest, β_0 is an intercept term, β_i is the coefficient associated with the corresponding explanatory variable x_i, $x=(1,x_1, x_2, \ldots, x_k)$ and $\beta=(\beta_0, \beta_1, \ldots, \beta_k)'$. So, in logistic regression, one estimates the log of the probability odds by a linear combination of the characteristic variables. Since $p/(1-p)$ takes values between 0 and ∞, $\log[p/(1-p)]$ takes values between $-\infty$ and $+\infty$. Taking exponentials on both sides of (1) leads to

$$p = \frac{\exp(x\beta)}{1 + \exp(x\beta)}.$$

Logistic regression can be used to obtain acceptance probability for FSCA data. In the simplest model let the applicant characteristics be $x=(x_1, \ldots, x_n)$ and let $o=(o_1, \ldots, o_m)$ be the offer characteristics. Then the basic logistic regression approach assume that the probability that an applicant with characteristics x will accept offer o satisfies

$$\log[p(1-p)] = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \beta_{n+1} o_1$$
$$+ \cdots + \beta_{n+m} o_m = y\beta, \quad (2)$$

where $y=(1, x_1, \ldots, x_n, o_1, \ldots, o_m)$ and $\beta=(\beta_0, \beta_1, \ldots, \beta_{n+m})'$.

An obvious extension is to allow interaction characteristics $i=(i_1, \ldots, i_p)$ which are combined offer and application characteristics, e.g. take value 1 if customer likes travel and offer gives no fee foreign exchange purchase; 0 otherwise.

The first step in building a credit scorecard using logistic regression on a sample of past customers is to coarse classifying the characteristics and we will do this for all versions of our acceptance scorecard. In general, the coarse classifying procedure splits the values of a continuous characteristic into bands and the values of a discrete characteristic into groups of values, where the values in each band or group tend to have roughly the same odds of the to outcomes in the original data sample. The binary variables corresponding to the indicator variable for each of the bands and groups chosen are then the ones used in the logistic regression. Coarse classifying improves the robustness of the scorecard being developed, since it increase the size of the group with a particular regression coefficient. More importantly for continuous variables, it allows for non-monotonicity between the characteristic values and the probability of the outcomes. There are some additional difficulties in using coarse classifying in this context since there are substantial differences in the offers being made to the different consumers in the data set. Ways of dealing with this were investigated in Jung and Thomas (2003) but here, we will use the standard approach (Thomas et al., 2004) where one starts with a very fine classification (every decile of a continuous variable and all values of a categorical variable say) and use the chi-square and information statistics to combine some of these classes.

The coarse classifying led to the following bands and groups being constructed

Age: 20 or less (Age 1); 21–30(Age 2); 31+(Age 3); missing value is the reference group (remember this is a student population)

Children: 1 or more; 0 is the reference group

Overdraft: £1250 or less; reference group is more than £1250

Current account interest: more than 1% (interest); reference group is 1% or less

Insurance; discount on any form of insurance except musical insturments (insurance); insurance on musical instrument (music); reference group is no discount on any insurance.

The logistic model is then built on the application and offer characteristics only using the 265 cases in the training sample. The offer variables which are significant in the regression and hence have a major impact in the acceptance scorecard were overdraft, interest and insurance. Table 2 gives the result of the logistic regression.

The classification power and robustness of the scorecard developed was tested by using the holdout sample and the results

Table 2
Result of logistic regression on offer and applicant characteristics

Variable	Coefficient	S.E.	Wald	DF	Sig	Exp (B)
Age1	1.163	0.757	2.359	1	0.125	3.199
Age2	1.341	0.685	3.829	1	0.050	3.822
Age3	0.484	0.810	357	1	0.550	1.623
Sex	0.775	0.460	2.833	1	0.092	2.170
Children	1.789	1.237	2.092	1	0.148	5.981
Cinema	−0.704	0.465	2.298	1	0.130	0.494
Diy	2.540	1.244	4.164	1	0.041	12.674
Overdraft	−2.465	0.846	8.483	1	0.004	0.085
Travelmoney	0.920	0.495	3.459	1	0.063	2.510
Insurance	1.579	0.845	3.494	1	0.062	4.850
Constant	−0.878	0.873	1.012	1	0.315	0.416

Table 3
Classification results using the logistic regression model

	Training data		Holdout data		Whole data	
	Actual numbers	LR	Actual numbers	LR	Actual numbers	LR
Y-predicted Y	155	121	39	29	194	150
Y-predicted N	0	34	0	10	0	44
N-predicted N	110	57	27	15	137	72
N-predicted Y	0	53	0	12	0	65

given in Table 3 where Y means the applicant said Yes to the offer they were given and N means that they said No, while predicted Y and predicted N is what the model predicted they would say.

Since, the sample is relatively small the analysis was repeated using the leave one out approach, where the regression is in turn built on all but one of the sample and tested on the remaining sample point. The results were very similar to those in Table 3.

In the previous scorecard, there were no applicant-offer interaction variables present. The model was extended by including such variables concentrating on interactions between the seven application variables and the three offer variables that appeared in the 'top ten' variables of the original scorecard. Chi-squared statistics, the F information statistic and D concordance statistics were used to identify the important interactions and three interactions were identified as relevant. However, when these were introduced into the scorecard there was only a minor improvement with one extra case being correctly classified. This is slightly disappointing in that it says the relevant ranking of the probability of accepting the different offers would be the same for all applicants. One suspects this is because of the limited number of applicants in the sample and the artificiality of the situation. In real situations, these interaction terms would be expected to play more of a role.

4. Exact cut-off on dominant characteristic using linear programming

The second approach to developing a model to determine which offer to make to an applicant, assumes that there is a dominant offer characteristic, where the attractiveness of the offer increases (or decreases) monotonically as this characteristic's value increases. The idea is that, given the other offer and applicant characteristics, one can identify the value of this dominant characteristic at which this applicant would accept the offer. One could then use this value in profit calculations to see whether it is profitable to make such an offer. In the credit card context, this dominant characteristic could be the interest rate charged for borrowers or the credit limit for transactors,

but for student bank accounts, where the overdraft is interest free, the overdraft limit is clearly the most important offer feature to most students.

So again assume applicant characteristics $x = (x_1, x_2, ..., x_n)$, offer characteristics $o = (o_2, ..., o_n)$ with $O = o_1$ being the dominant offer characteristic and the interaction characteristics i. We are interested in determining the accept/reject level of O, O^*, as a linear function of x, o and i. Hence, we assume

$$O^* = c_0 + c_1 x + c_2 o + c_3 i = cy, \qquad (3)$$

where $y = (x, o, i)$.

Taking a sample of previous applicants, if applicant i (with characteristics y^i) accepted an offer of o^i then $o^i \geq cy^i$ while if applicant j (with characteristics y^j) rejected on offer of o^j then $o^j \leq cy^j$, where we are assuming that the likelihood of acceptance increases as O increases. Hence, we can use linear programming to determine the coefficients c as follows.

Let the sample of previous customers be labeled 1 to $n(a) + n(r)$ where $i = 1, 2, ..., n(a)$ accepted the offer and $j = n(a) + 1, n(a) + 2, ..., n(a) + n(r)$ rejected the offer. Let applicant i have applicant/offer characteristics $y^i = (y_1^i, ..., y_p^i)$ and be made an offer o^i. Then to find the coefficients c that give the best estimate of the overdraft accept/reject indifference level we want to solve the following linear programme.

Minimise $\quad e_1 + ... + e_{n(a)+n(r)}$

$$\begin{aligned} \text{Subject to} \quad & o^i + e_i \geq c_1 y_1^i + c_2 y_2^i + \cdots + c_p y_p^i, \quad i = 1, 2, ..., n(a), \\ & o^i - e_j \leq c_1 y_1^j + c_2 y_2^j + \cdots + c_p y_p^j, \quad j = n(a) + 1, ..., n(a) + n(r), \\ & e_i \geq 0, \qquad\qquad\qquad\qquad\qquad i = 1, 2, ..., n(a) + n(r). \end{aligned} \qquad (4)$$

This has some relationship with the linear programming formulation of how to build a default/not default credit scorecard, but the introduction of different cut-off levels for each individual leads to a somewhat different formulation.

This model was applied to the fantasy account data set, using the application and offer variables and their characteristics identified in Section 3. All the offer and application variables were used and the coarse classifying developed in the previous section gave 32 binary variables to put into the linear programme, but because of the results in Section 3 no interaction variables were included.

The results were as follows (Table 4).

One could imagine that linear programming does particularly well here because of the number of variables compared with the size of the training sample. As the sample sizes

Table 4
Classification results using linear programming model

	Training data		Holdout data		Whole data	
	Actual numbers	LP	Actual numbers	LP	Actual numbers	LP
Y-predicted Y	155	138	39	33	194	171
Y-predicted N	0	17	0	6	0	23
N-predicted N	110	102	27	26	137	128
N-predicted Y	0	8	0	1	0	9

Table 5
Coefficients in Linear programming approach

Attribute	All age	0 Credit card	1 Credit card	2 Credit card	3+ Credit card	Musical instr. insurance
Score	450	650	0	400	900	−250
Attribute	All other or no insurance	Intro offer-CD player or miniTV £40	No into offer/brewing kit/£40	CDs and CD vouchers	Rail card	All other or no insurance
Score	0	550	150	650	700	0

increases it is likely the improvement in classification accuracy over the logistic regression approach may diminish a little.

The variables which had significant coefficients were given in Table 5.

Clearly, these values should not be taken as definitive given the fantasy nature of the data but they do show how this approach could lead to useful information as part of the scorecard building process. The scorecard in Table 5 says those with 1 credit card are the ones who will accept the lowest overdraft while those with 0 credit cards want an overdraft of £600 more before they accept, those with 2 credit cards want £400 more while the worldly wise with 3+ credit cards want £900 more on their overdraft limit before they accept. Similarly, offering insurance on musical instruments means students would accept a loan with £250 less on the overdraft but other forms of free insurance make no difference. The results suggest that initial offers are counterproductive with offering a mini TV means the overdraft limit being looked for goes up by £550 while offering railcards mean it goes up by £700. For a student aged 21 with no credit cards an offer which includes no insurance and free CDs would mean the credit limit needs to be $450 + 650 + 0 + 650 = £1750$ before the student will accept.

5. Cut off distribution on dominant characteristic using accelerated life approach

In the previous approach, we assumed there was an exact value of the dominant characteristic at which an applicant would change from rejecting the offer to accepting it. A weaker assumption would be to say there is a probability distribution over the values of the dominant characteristic of where the cut-off point between accepting and rejecting lies. Arguments for assuming a probability distribution of acceptance rather than an exact cut-off point as in the previous section is that the changes in the individual's desires, economic circumstances and the environment in which the offer is made, all of which can fluctuate rapidly, could mean the same person makes different decisions to the same offer at different times. It might also reflect an applicant's inability to make decisive judgments between incrementally different offers. In this section we show how using accelerated life model, which is one of the models used in survival analysis, can help one determine such a distribution. Survival analysis is the area of statistics that deals with analysis of lifetime data, but recently (Narain, 2004; Thomas, Banasik, & Crook, 1999; Stepanova & Thomas, 2002), one has been able to use survival analysis approaches, especially the proportional hazards model and the accelerated life models to build consumer credit assessment systems. Here, we present a new application of accelerated life models in the consumer acceptance estimation area.

To apply the accelerated life model for FSCA data, we again assume that there is an important offer characteristic and the probability of accepting the offer increases or decreases monotonically as this dominant offer characteristic's value increases. The idea is that given the other offer and applicant characteristics one can identify the value of this dominant characteristic, or at least the distribution function of the value, at which the applicant would accept the offer.

To estimate this distribution function assume O_1 is the dominant offer characteristic. If one has applicant characteristics x and offer characteristic (t, o) where t is the value of the dominant monotone characteristic O_1, then we are interested in the probability of an applicant with characteristics x accepting offer (t, o). Thus, if T is the lowest value of O_1 at which the offer will be acceptable, then

Prob{individual with characteristic x accepts offer (t, o)} = Prob$\{T \leq t | y = (x, o)\} = F(t | y = (x, o))$

where $y = (x_1, ..., x_l, o_1, ... o_m)$ and hence that the probability of an individual with characteristic x rejecting offer (t, o) is given by

$$S(t | y = (x, o)) = 1 - F((t | y = (x, o)).$$

This is known in survival analysis as the survival function. The accelerated life model can be applied to estimate the reject probability for FSCA Data where Table 1 shows the applicant and offer characteristic used in analysis and the overdraft limit is taken

as the dominant characteristic. Notice that all the data will be either right or left censored. If applicant i with characteristics x_i accepts offer (t, o_i) then all we can say is $T \leq t$. It applicant j with characteristics x_j rejects offer (t, o_j) then all we can say is $T \geq t$. Thus, we never observe uncensored data. In the accelerated life model, one define the survival (reject) function by

$$S(t|y) = S_0(e^{y\beta}t), \tag{5}$$

where $y = (1, x_1, \ldots, x_l, o_1, \ldots, o_m)$, $\beta = (1, \beta_0, \beta_1, \ldots, \beta_{l+m})'$ and $S_0(\cdot)$ is a baseline survival function. One common accelerated life model in survival analysis is to take $S_0(\cdot)$ to be the Weibull distribution. This gives a baseline survival function of

$$S_0(t) = \exp\{-(\lambda t)^k\}, \tag{6}$$

where λ and k are the scale and shape parameters of the Weibull distribution. From (5) and (6), the survival function in accelerated life model can be expressed as follows.

$$S(t|y) = \exp\{-(\lambda \exp(b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_l x_l$$
$$+ b_{l+1} o_l + \cdots + b_{l+m} o_m)t)^k\}$$
$$= \exp\{-(\lambda \exp(y\beta)t)^k\}. \tag{7}$$

By applying the accelerated life model of doubly censored data, we can obtain the likelihood function as follows.

$$L(\theta) = \prod_{i=1}^{n(a)} (1 - S(t_i|y_i)) \prod_{j=n(a)+1}^{n(a)+n(r)} S(t_j|y_j)$$
$$= \prod_{i=1}^{n(a)} [1 - \exp\{-(\lambda \exp(y_i\beta)t_i)^k\}]$$
$$\times \prod_{j=n(a)+1}^{n(a)+n(r)} \exp\{-(\lambda \exp(y_j\beta)t_j)^k\}.$$

The maximum likelihood estimates of λ, k and β in the Weibull based accelerated life model are then obtained using Newton–Raphson methods.

It is still necessary to coarse classify the variables used but we can use the accelerated life approach not just to build the model but also to do the coarse classifying. This overcomes the problem that in the traditional coarse classifying approach we used in the logistic regression and linear programming methods described earlier we do not consider the actual offer made even though there is a strong interaction between the offer level and the accept–reject decision. The coarse classifying method used in this AL model follows closely the approach used for applying proportional hazards models in credit scoring (Stepanova & Thomas, 2002). It consists of the following steps

Step 1. (Continuous characteristic) Split the characteristic into n binary variables with approximately equal number of observations in each variable. (Discrete characteristic) A binary variable is created for each attribute of the characteristic.
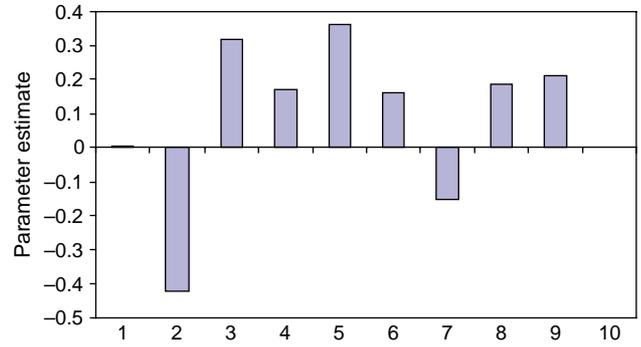
Fig. 1. AL parameter estimates for introductory gift characteristic.

Step 2. Apply AL model with these binary variables.
Step 3. Chart parameter estimates.
Step 4. Choose the splits based on similarity of parameter estimates.

As an illustration of coarse classifying using the AL model, consider the introductory gift characteristic. Fig. 1 shows the histogram of the parameter estimates and consideration of this leads to the classification given in Table 6.

Using the binary variables obtained by this coarse classifying the Accelerated life model is built on the same training sample as previously and tested on the holdout sample. The coefficients are estimated using (7) and the results give an accelerated life model of the form

$$S(t|y) = \exp\{-(\lambda \exp(\beta_0 + \beta_1 \text{Var}_1 + \cdots + \beta_q \text{Var}_q)t)^k\},$$

where Var_1, Var_2, …, Var_q are the indicator variables obtained by the coarse classifying, β_0, β_1, …, β_q are the parameters to be estimated and t is the overdraft likely to be accepted by an applicant with application and other offer characteristics y.

To compare the AL model with the LR and the LP models, all the variables are included in the AL model. Table 7 shows the results on a training, holdout and whole sample for all three methods.

Thus, on this data set the accelerated life model classifies the least well but is close to the results for the logistic regression approach. Part of this is due to the smallness of the set compared with the number of variables available and part may be due to the fact there were only a few overdraft levels being offered. If there were only one level of overdraft offered, then estimating a distribution of what level an applicant would accept must be less precise than estimating if they would accept the offer at that given level which in a sense is what the logistic regression is strongest at

Table 6
Coarse classifying of the introductory gift characteristic

Binary variable	Type of offer	No. of observations	
		Accept	Reject
Introductory_1	1, 10	36	45
Introductory_2	2, 7	107	47
Introductory_3	3, 4, 5, 6, 8, 9	53	45

Table 7
Classification results using AR and the other models

	Training data				Holdout data				Whole data			
	Actuals	LR	LP	AL	Actuals	LR	LP	AL	Actuals	LR	LP	AL
Y predicted Y	155	121	138	124	39	29	33	28	194	150	171	152
Y predicted N	0	34	17	31	0	10	6	11	0	44	23	42
N predicted N	110	57	102	50	27	15	22	26	137	72	128	62
N predicted Y	0	53	8	60	0	12	5	1	0	65	9	75

6. Conclusions

This paper introduces three techniques which can be used to build models of the probabilities that a particular consumer will accept different variants of a generic borrowing product like a credit card or account with an overdraft facility. It derives a data set based on students acceptance or rejection of a fantasy student account offer. Thus, one must include many of the caveats when one uses data which is essentially obtained from gaming experiments rather than from real experience. However, the paper does show that it is possible to build acceptance probability models using such data and makes a preliminary investigation of three different approaches. We believe that two of these—linear programming and accelerated life models— have not been tried before in this context. All three approaches are technically feasible and can result in real time decisions about which variant of the product to offer the current applicant in order to maximize profit, though two of them do require the notion of a dominant offer characteristic. For many products, it does seem reasonable to assume that one characteristic has the necessary monotone properties to apply such procedures.

Although the results suggest the linear programming approach to estimate a cut-off level on the dominant characteristic does clearly make the best predictions on this data set both on the training sample and on the holdout sample, we feel these results must be taken with a great deal of caution because of the relatively large number of binary variables available compared with the size of the sample.

We believe these probability acceptance models will become increasingly important as the consumer lending market matures and it becomes a buyers rather than a sellers market. They are ideally suited to the interactive application processes that modern telecommunication technology is supporting. They also satisfy the customer relationship marketing credo of tailoring the product to the customer.

Acknowledgements

References

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7, 179–188.

Freed, N., & Glover, F. (1981). A linear programming approach to the discriminant problem. *Decision Sciences*, 12, 68–74.

Geoffrion, A. M., & Krishnan, R. (2003). E-business and management science: Mutual impacts. *Management Science*, 49, 1275–1286.

Hand, D. J. (1981). *Discrimination and classification*. Chichester, UK: Wiley.

Jung, K. M., & Thomas, L. C. (2004). A note on coarse classifying in acceptance scorecards, Working paper M04-16, School of Management, University of Southampton.

Kannan, P. K., & Raghav Rao, H. (2001). Decision support issues in customer relationship management and interactive marketing for e-commerce. *Decision Support Systems*, 32, 83–94.

Karuga, G. G., Khraban, A. M., Nair, S. K., & Rice, D. O. (2001). AdPalette: An algorithm for customizing online advertisements on the fly. *Decision Support Systems*, 32, 85–106.

Mangasarian, O. L. (1965). Linear and nonlinear separation of patterns by linear programming. *Operations Research*, 13, 444–452.

Montgomery, A. L. (2001). Applying quantitative marketing techniques to the Internet. *Interfaces*, 31, 90–108.

Murthi, B. P. S., & Sarkar, S. (2003). The role of the management sciences in research on personalization. *Management Science*, 49, 1344–1362.

Narain, B. (2004). Survival analysis and the credit granting decision. In L. C. Thomas, J. N. Crook, & D. B. Edelman (Eds.), *Readings in credit scoring* (pp. 235–245). Oxford: Oxford University Press (Reprinted).

Raghu, T. S., Kannan, P. K., Rao, H. R., & Whinstom, A. B. (2001). Dynamic profiling of consumers for customized offerings over the Internet; a model and analysis. *Decision Support Systems*, 32, 117–134.

Rosenberg, E., & Gleit, A. (1999). Quantitative methods in credit management: A survey. *Operations Research*, 42, 589–613.

Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The value of purchase history data in target marketing. *Marketing Science*, 15, 321–340.

Stepanova, M., & Thomas, L. C. (2002). Survival analysis method for personal loan data. *Operations Research*, 50, 277–289.

Thomas, L. C. (2000). A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16, 149–172.

Thomas, L. C., Banasik, J., & Crook, J. N. (1999). Not if but when loans default. *Journal of the Operational Research Society*, 50, 1185–1190.

Thomas, L. C., Edelman, D. B., & Crook, J. N. (2002). *Credit scoring and its applications*. Philadelphia, PA: SIAM.

Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behaviour. *Journal of Financial and Quantitative Analysis*, 15, 757–770.