

ESTIMATION OF A CREDIT SCORING MODEL FOR LENDERS COMPANY

Felipe Alonso Arias
Tutor: Juan Sebastián Bravo
Co-tutor: Francisco Iván Zuluaga

Progress Presentation Research Practise 3

Mathematical Sciences Department, EAFIT University
Linkvest Capital LLC, Miami, Fl
September 25 2015

Remember...

- ▶ Mortgage broker
- ▶ People ask for a loan
- ▶ Linkvest analyze the background
- ▶ Make decisions based on experience
- ▶ Lend money/high interest rate
- ▶ Probability of default



The model

The logistic regression model [Cantón et al., 2010] can be formulated as

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \mu$$

where μ is the disturbance and p is the probability of default [Hand & Henley, 1997] and can be estimated as follows

$$p = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}} \quad (1)$$

The model

Then, using (1) our response is going to be

$$Default = \log \left(\frac{p}{1-p} \right)$$

This response is going to give us a number near to 0 or 1, 0 for no paying and 1 for paying [Orgler, 1970].

Three big problems

- ▶ Incomplete loan applications
 - ▶ We are now using loan application for all loans
- ▶ Decisions based on previous situations
 - ▶ Now we will use the past applications, not the experience
- ▶ How to know if the person is going to pay?
 - ▶ Credit score model using information retrieved with previous loans

General objective

- ▶ Estimate a credit score model to improve lending decisions on a person using historical data
 - ▶ Working on it...

Specific objectives

- ▶ Analyze the best model to improve in our case based in the data
 - ▶ Logistic regression model [Cantón et al., 2010]
- ▶ Identify the most influential variables for the company when lending money [Altman, 1968]
 - ▶ Already defined
- ▶ Specify the model in mathematical terms
 - ▶ Already defined
- ▶ Estimate the probabilities of default
- ▶ Analyze the results for decision making

Variables

- ▶ Initially 47 different variables [Cantón et al., 2010]
- ▶ 28 defined as relevant variables [Altman, 1968]
- ▶ Amount/interest rate/months
- ▶ Age/years school/married/dependents/own home/ethnicity/sex
- ▶ Self-employed/monthly income/monthly expense/cash or assets/fixed assets
- ▶ Judgment/bankrupt/foreclosure/lawsuit/obligated on a loan/federal debt/child support/borrowed/co-maker/citizen/resident/PR/ownership

Mathematical model

$$Z = \Sigma\Psi$$

where

$$\Sigma = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_{28}] \quad \Psi = \begin{bmatrix} 1 \\ V_1 \\ V_2 \\ \vdots \\ V_{27} \end{bmatrix}$$

Variables

V_1 Amount

V_2 Interest

V_3 Months

V_4 Age

V_5 Years school

V_6 Married

V_7 Dependents

V_8 Home

V_9 Employment

V_{10} Income

V_{11} Expense

V_{12} Cash

V_{13} Assets

V_{14} Questions
($V_{14}, V_{15}, \dots, V_{25}$)

V_{26} Ethnicity

V_{27} Gender

Schedule

Week	Task
4-5	Proposal report✓
6-7	Specification of the model✓
8-9	Estimation of the model
10	Progress presentation✓
11-14	More estimations
15-16	Final report
17	Extra improvements
18	Final presentation

Resources

- Altman, E. I. (1968, September). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, XXIII(4), 589-609.
- Cantón, S. R., Rubio, J. L., & Blasco, D. C. (2010, June). A Credit Scoring Model for Institutions of Microfinance under the Basel II Normative. *Journal of Economics, Finance and Administrative Science*, 15(28).
- Hand, D. J., & Henley, W. E. (1997). Statistical Classification Methods in Customer Credit Scoring: A review. *Journal of the Royal Statistical Association*, 160(Part 3), 523-541.
- Orgler, Y. E. (1970, November). A Credit Scoring Model for Commercial Loans. *Journal of Money, Credit and Banking*, 2(4), 435-445.

Thanks for attending!