

Contents

List of Figures	vii
List of Tables	ix
Glossary	xi
1 Introduction	1
1.1 Uses of internal ratings	1
1.2 Outline of analysis	3
2 2-group classification versus rating of firms	7
2.1 Introduction	7
2.2 Architecture of internal rating systems	9
2.2.1 Risk factors	12
2.2.2 Value functions	19
2.2.3 Risk factor weights	21
2.2.4 Rating assignment rule	22
2.3 Statistical credit scoring	24
2.3.1 Linear discriminant analysis	24
2.3.2 The linear probability model	27
2.3.3 Logistic regression	29
2.3.4 The logistic perceptron	31
2.4 Comparison of estimation methods	33
2.4.1 Comparison of CML and LS-BP	33
2.4.2 Comparison of CML and LDA	34
2.5 A location model for combinations of risk factors	36
2.6 Monte Carlo setup	39
2.6.1 Specification of the qualitative risk factors	39
2.6.2 Specification of the quantitative risk factors	40
2.6.3 Measures of classification and rating assignment accuracy	44

2.7	Results	48
2.7.1	Accuracy of 2-group classifications	50
2.7.2	Accuracy of rating assignments	53
2.8	Concluding remarks	57
3	Estimation of transition probabilities	61
3.1	Introduction	61
3.2	Literature review	63
3.3	Methods for estimating transition intensities	64
3.3.1	The continuous-time case	65
3.3.2	The discrete-time case	67
3.3.3	Comparison of estimation methods	69
3.4	External versus internal rating data	69
3.4.1	The temporal difference	71
3.4.2	The monitoring intensity-based difference	72
3.4.3	The internal incentive-based difference	74
3.5	Implications for transition probability estimation	75
3.5.1	Implications from the temporal difference	75
3.5.2	Implications from the monitoring intensity-based difference	78
3.6	A Maximum Likelihood estimation method	79
3.7	Empirical application	81
3.7.1	The data	81
3.7.2	Results	84
3.8	Concluding remarks	103
4	Missing ratings	105
4.1	Introduction	105
4.2	Inference under missing rating data	107
4.2.1	Agency problems as driving forces for missing ratings	108
4.2.2	Inference without assumptions concerning the missingness mechanism	113
4.2.3	Inference when internal ratings are missing completely at random	114
4.2.4	Inference when internal ratings are missing at random	116
4.2.5	Inference when internal ratings are non-ignorable missing	119
4.3	Empirical application	124
4.3.1	The data and variables	124
4.3.2	Results	130
4.4	Estimation of non-default transition probabilities	140
4.5	Concluding remarks	144

5 Conclusion	147
A The EM algorithm	151
A.1 The case of non-ignorable missingness	151
A.2 The case of data missing at random	156
B Additional tables	159
References	165

List of Figures

2.1	A possible transformation curve	22
2.2	The logistic perceptron	31
2.3	2-group classification versus assignment of ordinal rating grades	35
2.4	Course of the difference (<i>dif</i>) between the estimated unconditional error rate $\hat{e}_{(0.632)}$ and the optimal error rate e_{LM} for the logistic regression (CML) and the discriminant analysis (LDA) approach of estimating scoring model A	51
2.5	Course of the difference (<i>dif</i>) between the estimated unconditional error rate $\hat{e}_{(0.632)}$ and the optimal error rate e_{LM} for the logistic regression (CML) and the discriminant analysis (LDA) approach of estimating scoring model B	52
2.6	Course of the difference (<i>dif</i>) between the estimated unconditional error rate $\hat{e}_{(0.632)}$ and the optimal error rate e_{LM} for the two discriminant analysis approaches (LDA and FL/LDA) of estimating scoring model A	53
2.7	Course of the difference (<i>dif</i>) between the estimated unconditional error rate $\hat{e}_{(0.632)}$ and the optimal error rate e_{LM} for the two discriminant analysis approaches (LDA and FL/LDA) of estimating scoring model B	54
2.8	Course of the average proportion of firms with correspondence between true and estimated rating assignments for the three approaches (CML, LDA and FL/LDA) of estimating scoring model A	55
2.9	Course of the average proportion of firms with correspondence between true and estimated rating assignments for the three approaches (CML, LDA and FL/LDA) of estimating scoring model B	56

2.10	Course of the average proportion of firms with a small (≤ 0.005) absolute difference among true and estimated capital requirements for the three approaches (CML, LDA and FL/LDA) of estimating scoring model A	58
2.11	Course of the average proportion of firms with a small (≤ 0.005) absolute difference among true and estimated capital requirements for the three approaches (CML, LDA and FL/LDA) of estimating scoring model B	59
3.1	Evolution of credit quality for a fictitious borrower over 66 months	73
4.1	Tables with complete and incomplete data	117
4.2	Estimated cumulative hazard functions	131

List of Tables

2.1	Criteria catalogues of major rating agencies and a representative bank	13
2.2	International survey of financial ratios used in credit scoring models	14
2.3	Group distance	41
2.4	Features of data samples used by Moody's	42
2.5	Parameter constellation for the 27 distinct simulation situations	43
2.6	Theoretical error rates	49
3.1	Different data situations and corresponding ML procedures for estimating the generator of a homogeneous continuous-time Markov process.	70
3.2	Risk/size distribution (1996-2002)	82
3.3	Rating distribution over time	83
3.4	Definitions of variables and hypotheses	86
3.5	Determinants of interexamination time	88
3.6	The ML estimator of the generator Q based upon discrete-time observations for the two periods 1996-2000 and 1996-2002	92
3.7	Transition probabilities: comparing estimation methods	93
3.8	Internal rating prediction model	95
3.9	Bootstrap results for the distance metric $\Delta M_{SVD}(\hat{P}_{ML}, \hat{P}_C)$ between ML and cohort method	97
3.10	Characteristics of sample portfolios and calibration of portfolio model	98
3.11	Credit risk capital: comparing estimation methods	100
3.12	Results for the test of business cycle effects and non-Markov behavior of transition intensities to a neighboring rating class	102
4.1	Different default probability estimators	125
4.2	Size distribution (all data)	126
4.3	Risk distribution (1996-2002, completely observed data)	127

4.4	Distribution of default events (1996-2002)	128
4.5	Rating class default probability estimates	133
4.6	The default model	134
4.7	Bootstrap results for the difference in estimates	136
4.8	The missingness model	137
4.9	The rating distribution	138
B.1	Mean and standard deviation of area under the ROC curve (AUC) for scoring model A and $n=1,000$	160
B.2	Mean and standard deviation of area under the ROC curve (AUC) for scoring model A and $n=10,000$	161
B.3	Mean and standard deviation of area under the ROC curve (AUC) for scoring model B and $n=1,000$	162
B.4	Mean and standard deviation of area under the ROC curve (AUC) for scoring model B and $n=10,000$	163
B.5	Transition probabilities: comparing estimation methods	164