

Bernhard Christian Kübler

Risk Classification by Means of Clustering

**Schriften zum Controlling,
Finanz- und Risikomanagement**

Herausgegeben von Andreas Brieden, Thomas Hartung,
Bernhard Hirsch und Andreas Schüler



PETER LANG

Internationaler Verlag der Wissenschaften

Contents

List of Figures	7
List of Abbreviations	8
1 Introduction	11
1.1 Exploratory Focus	11
1.2 Economic Incentive	16
1.2.1 On the Relevance of Classification in Eco- nomics	17
1.2.2 Application Areas of Cluster Analysis . .	19
1.3 Notation	24
1.4 A Plan of this Investigation	29
2 Fundamentals of Premium Pricing	32
2.1 General Remarks	32
2.1.1 Objectives and Techniques of Premium Pric- ing	33
2.1.2 Determinants of Preferable Classification Systems	43

2.2	Loss Models	44
2.2.1	The Individual Model	44
2.2.2	The Collective Model	47
2.3	Optimal Estimation and Credibility Theory . . .	50
2.3.1	Limited Fluctuation Theory	51
2.3.2	Greatest Accuracy Theory – a Bayesian Approach	54
3	Clustering in Actuarial Mathematics	58
3.1	Motivation	58
3.2	Distance and Similarity	64
3.2.1	Object Distance	64
3.2.2	Cluster Distance	70
3.2.3	Cluster Inhomogeneity	76
3.3	Cluster Criteria	78
3.4	Clustering Algorithms	82
3.4.1	Heuristic Algorithms	82
3.4.2	A Quadratic Optimization Model	84
3.5	Actuarial Problems from Cluster Analysis	86
3.5.1	Assessing Clustering Solutions	87
3.5.2	Selection of Tariff Variables	90
3.5.3	The Number-of-Classes-Problem	96
3.5.4	Determining Cluster Cardinalities in Ac- tuarial Applications	100

4	Cross-Classification	111
4.1	Proceeding	111
4.2	Premium Pricing and CC	114
4.2.1	The Method of Bailey and Simon	117
4.2.2	The Method of Marginal Totals	118
4.2.3	The Method of Least Squares	119
4.2.4	The Method of "Marginal Averages"	119
4.2.5	Direct Method	120
4.2.6	Generalized Linear Models	120
4.3	Cross-Classification and Cluster Analysis	122
4.3.1	Territory Classes	123
4.3.2	Type of Vehicle	127
4.3.3	Contiguous Clustering Problems	128
5	A New Classification System	132
5.1	Consequences of Our Previous Analysis	132
5.2	Problems Associated with CC	133
5.3	The MC	135
6	Empirical Investigation	142
6.1	Organization of the Data	142
6.1.1	Description of the Data Sets	142
6.1.2	Construction of a Suitable Collective	145
6.2	Risk Analysis	148
6.3	Application of MC	153

6.4	Application of CC	156
6.5	Economic Evaluation	161
6.5.1	Criteria	161
6.5.2	Calculations	164
6.6	Discussion and Outlook	166
A	Data	169
A.1	Random Number Generator	169
A.2	Claim Data	170
A.3	The Collective	171
	Bibliography	182