THE JOURNAL OF
Risk and Insurance

September 1995 • Volume 62, Nº 3

PUBLISHED BY THE AMERICAN RISK AND INSURANCE ASSOCIATION, INC.

INVITED ARTICLE

SYMPOSIUM ON INSURANCE SOLVENCY AND FINANCE

NOTES AND COMMUNICATIONS

© *The Journal of Risk and Insurance*, 1995 Vol. 62, No. 3, 447-482

Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification

Richard A. Derrig
Krzysztof M. Ostaszewski

ABSTRACT

Applications of fuzzy set theory to property-liability and life insurance have emerged in the last few years through the work of Lemaire (1990), Cummins and Derrig (1993, 1994), and Ostaszewski (1993). This article continues that line of research by providing an overview of fuzzy pattern recognition techniques and using them in clustering for risk and claims classification. The classic clustering problem of grouping towns into rating territories (DuMouchel, 1983; Conger, 1987) is revisited using these fuzzy methods and 1987 through 1990 Massachusetts automobile insurance data. The new problem of classifying claims in terms of suspected fraud is also addressed using these same fuzzy methods and data drawn from a study of 1989 bodily injury liability claims in Massachusetts.

Introduction

In 1961, Ellsberg presented the following paradox. An experiment was designed with two urns, each containing 100 balls, of which the first one was known to contain 50 red balls and 50 black balls, while no further information was given about the contents of the other urn. If asked to bet on the color of a ball drawn from one of the urns, most people were found indifferent as to which color they would choose no matter whether the ball was drawn from the first or the second urn. On the other hand, Ellsberg found that if people were asked which urn they would prefer to use for betting on either color, they

---

Richard A. Derrig is Senior Vice President of the Automobile Insurers Bureau of Massachusetts and Vice President–Research for the Insurance Fraud Bureau of Massachusetts. Krzysztof M. Ostaszewski is Associate Professor of Mathematics and Actuarial Program Director at the University of Louisville.

Krzysztof Ostaszewski has worked on this project at the University of Louisville with financial support from the Actuarial Education and Research Fund, and this support from AERF is gratefully acknowledged. The authors thank Jeff Strong and Robert Roesch of the Automobile Insurers Bureau for invaluable help in programming and performing calculations involved in this project, Herbert I. Weisberg for suggesting the fuzzy clustering of fraud assessment data, Ruy Cardoso for helpful comments on an early draft, Julie Jannuzzi for production of the document, and one anonymous reviewer.

consistently favored the first urn (no matter what color they were asked to bet on).

What seems to be present in this experiment is the participants' perception of uncertainty. When we say "uncertainty," the usual association is with "probability." The Ellsberg paradox illustrates that some other form of uncertainty can indeed exist. Probability theory provides no basis for the outcome of the Ellsberg experiment.

Klir and Folger (1988) analyze the semantic context of the term "uncertain" and arrive at the conclusion that there are two main types of uncertainty, captured by the terms "vagueness" and "ambiguity." Vagueness is associated with the difficulty of making sharp or precise distinctions among objects. "Ambiguity" is caused by situations where the choice between two or more alternatives is unspecified. The basic set of axioms of probability theory originating from Kolmogorov, rests on the assumption that the outcome of a random event can be observed and identified with precision. Any vagueness of observation is considered negligible, or not significant to the construction of the theoretical model. Yet one cannot escape the conclusion that forms of uncertainty represented by vagueness of observations, human perceptions, and interpretations, are missing from probabilistic models, which equate uncertainty with randomness (i.e., a sophisticated version of ambiguity).

Several reasons may exist for wanting to search for models of a form of uncertainty other than randomness. One is that vagueness is unavoidable. Given imprecision of natural language, or human perception of the phenomena observed, vagueness becomes a major factor in any attempt to model or predict the course of events. But there is more. When the phenomena observed become so complex that exact measurement involving all features considered significant would be impossible, or longer than economically feasible for study, mathematical precision is often abandoned in favor of more workable simple, but vague, "common sense" models. Thus, complexity of the problem may be another cause of vagueness.

These reasons were the driving force behind the development of the fuzzy set theory (FST). This field of applied mathematics has become a dynamic research and applications field, with success stories ranging from a fuzzy logic rice cooker to an artificial intelligence in control of Japan's Sendai subway system. The main idea of fuzzy set theory is to propose a model of uncertainty different from that given by probability, precisely because a different form of uncertainty is being modeled.

Fuzzy set theory was created in Zadeh's (1965) historic article. To present this basic idea, recall that a *characteristic function* of a subset E of a universe of discourse U is defined as

$$\chi_E(x) = \begin{cases} 1 \text{ if } x \in E \\ 0 \text{ if } x \notin E. \end{cases}$$

In other words, the characteristic function describes the membership of an element x in a set E. It equals one if x is a member of E, and zero otherwise.

Zadeh challenged the idea that membership in all sets behaves in the manner described above. One example would be the set of "tall people." We consistently talk about the set of "tall people," yet understand that the concept used is not precise. A person who is 5'11" is tall only to a certain degree, and yet such a person is not "not tall." Zadeh writes,

> The notion of fuzzy set provides a convenient point of departure for the construction of a conceptual framework which parallels in many respects the framework used in the case of ordinary sets, but is more general than the latter and, potentially, may prove to have a much wider scope of applicability, particularly in the fields of pattern classification and information processing. Essentially, such a framework provides a natural way of dealing with problems in which the source of imprecision is the absence of sharply defined criteria of class membership rather than the presence of random variables.

In the fuzzy set theory, membership of an element in a set is described by the *membership function* of the set. If U is the universe of discourse, and E is a fuzzy subset of U, the membership function μ_E:U→[0,1] assigns to every element x in the set Ẽ its degree of membership $\mu_E(x)$. We write either (E,μ_E) or E~ for that fuzzy set, to distinguish from the standard set notation E. The membership function is a generalization of the characteristic function of an ordinary set. Ordinary sets are termed *crisp sets* in fuzzy sets theory. They are considered a special case—a fuzzy set is crisp if, and only if, its membership function does not have fractional values.

On the basis of this definition, one then develops such concepts as set theoretic operations on fuzzy sets (union, intersection, etc.), as well as the notions of fuzzy numbers, fuzzy relations, fuzzy arithmetic, and approximate reasoning (known popularly as "fuzzy logic"). Pattern recognition, or the search for structure in data, provided the early impetus for developing FST because of the fundamental involvement of human perception (Dubois and Prade, 1980) and the inadequacy of standard mathematics to deal with complex and ill-defined systems (Bezdek and Pal, 1992). The formal development began with Zadeh (1965) introducing the principal concepts of FST. A complete presentation of FST is provided in Zimmerman (1991).

The first recognition of FST applicability to the problem of insurance underwriting is due to DeWit (1982). Lemaire (1990) sets out a more extensive agenda for FST in insurance theory, most notably in the financial aspects of the business. Under the auspices of the Society of Actuaries, Ostaszewski (1993) assembled a large number of possible applications of fuzzy set theory in actuarial science. His presentation includes such areas as economics of risk, time value of money, individual and collective models or risk, classification, assumptions, conservatism, and adjustment. Cummins and Derrig (1993, 1994) complement that work by exploring applications of fuzzy sets to property-liability insurance forecasting and pricing problems.

Here, we present a method of fuzzy pattern recognition for risk and claims classification. We apply fuzzy pattern recognition to two problems in Massachusetts private passenger automobile insurance: defining rating territories and classifying claims with regard to their suspected fraud content. Dubois and

Prade (1980), Bezdek (1981), and Kandel (1982) provide overviews of fuzzy techniques in pattern recognition. Zimmerman (1991) and Bezdek and Pal (1992) provide other valuable references on the subject.

The concept of a fuzzy set and the mathematical algorithms needed to implement classification using fuzzy techniques is described in the next section. Grouping towns in Massachusetts into rating territories for risk classification purposes is viewed as a fuzzy clustering problem because many towns can be strongly related to two or more territories, thereby creating a border problem: to which of several related territories should a town be assigned. We also explore the influence of geographical proximity on the resulting fuzzy territories and classification of claims by their suspected fraudulent content. A final section summarizes and provides some alternative and future directions for FST in risk and claims classification problems.

Algorithms for Fuzzy Classification

Lemaire (1990) and Ostaszewski (1993) point out that insurance risk classification often resorts either to vague methods—as in the case of using multiple ill-defined personal criteria to identify good risks to underwrite—or methods that are excessively precise—as in the case of a person who fails to classify as a preferred risk for life insurance application because his or her body weight exceeds the stated limit by half a pound. Kandel (1982), writing from a different perspective, says: "In a very fundamental way, the intimate relation between the theory of fuzzy sets and the theory of pattern recognition and classification rests on the fact that most real-world classes are fuzzy in nature." This is exactly the reason that we propose to utilize the methodology of fuzzy clustering in territorial classification and to extend that method to classifying claims for suspected fraud.

Kandel (1982) classifies various techniques of fuzzy pattern recognition. *Syntactic* techniques apply when the pattern sought is related to the formal structure of the language. *Semantic* techniques apply to those producing fuzzy partitions of data sets. According to Bezdek and Pal (1992), the first choice faced by a pattern recognition system designer is that of process description. The designer may choose from among syntactic, numerical, contextual, rule-based, hybrid, and fuzzy process descriptions. Feature analysis is the next design step, in which data (generally given in the form of a data vector containing information about the analyzed objects) may be subjected to preprocessing, displays, and extraction. Next, semantic clustering algorithms, generating actual structures in data, are identified. Finally, the designer addresses cluster validity and optimality.

We use a fuzzy pattern recognition technique given by Bezdek (1981). In the classification of Bezdek and Pal (1992), it can be described as a numerical process description, fuzzy c-means iterative semantic algorithm. Because the data we analyze are in the form of numerical vectors (i.e., vectors in a euclidean space), with a number of clusters sought predetermined, we consider the

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.