Volume 20, No. 1 April 1990

ASTIN BULLETIN

A Journal of the International Actuarial Association

CONTENTS

Ceuterick

EDITORIAL POLICY

ASTIN BULLETIN started in 1958 as a journal providing an outlet for actuarial studies in non-life insurance Since then a well-established non-life methodology has resulted, which is also applicable to other fields of insurance For that reason *ASTIN BULLETIN* will publish papers written from any quantitative point of view -- whether actuarial, econometric, engineering, mathematical, statistical, etc —attacking theoretical and applied problems in any field faced with elements of insurance and risk

ASTIN BULLETIN appears twice a year, each issue consisting of about 80 pages

Details concerning submission of manuscripts are given on the inside back cover

MEMBERSHIP

ASTIN is a section of the International Actuarial Association (IAA) Membership is open automatically to all IAA members and under certain conditions to non-members also Applications for membership can be made through the National Correspondent or, in the case of countries not represented by a national correspondent, through a member of the Committee of ASTIN

Members receive *ASTIN BULLETIN* free of charge

SUBSCRIPTION AND BACK ISSUES

ASTIN BULLETIN is published and printed for ASTIN by Ceuterick s a , Brusselsestraat 153, B-3000 Leuven, Belgium

All queries and communications concerning subscriptions, including claims and address changes, and concerning back issues should be sent to Ceuterick

The current subscription or back issue price per volume of 2 issues including postage is BEF 1700

INDEX TO VOLUMES 1-13

The Cumulative Index to Volumes 1-13 is also published for ASTIN by Ceuterick at the above address and is available for the price of BEF 340

EDITORIAL AND ANNOUNCEMENTS

GUEST EDITORIAL

SOLVENCY CONTROL OF INSURERS — A CHALLENGE TO ACTUARIAL SCIENCE

Solvency of insurers is a highlight of actuarial study in our time. The topic is regularly discussed in the actuarial literature and at actuarial conferences. Even monographs and special meetings are entirely devoted to it, and a number of working parties—national as well as international—have been commissioned to work out practical solvency requirements and routines of solvency control Some general reasons for the prominence of the topic are obvious: to an insurance company, like any other business, prevention of negative results is of vital importance — preferably profit should be produced, and supervisory authorities conducting public affairs must ascertain that the insurers are maintaining their part of the social security system.

Special reasons for the evergrowing prominence of the topic nowadays are to be found in the rapid changes in the market Modern economic life is characterized by the emergence of progressively bigger decision making bodies — firms and organizations. In particular, their role as purchasers of insurance is quite different from that of yesterday's typically smaller decision units: they have the capacity for selfinsurance by e g. captives or pension funds or simply by not buying insurance, they often possess know-how in risk assessment; and being buyers of insurance on a large scale, they are able to compare premium expenses to benefits and thereby judge the fairness of the prices of insurance products. These changes on the demand side have enforced increased competition between insurers. The globalization of the insurance business pulls in the same direction. In their struggle for shares in a competitive market, the insurers launch myriads of new products designed for progressively more specific—hence smaller—groups of risks, and they quote premiums close to, and sometimes even below the net premium. It is a dilemma that the need for more accurate risk assessment is accompanied by a deterioration of statistical databases With the dissolution of the former cartel-like cooperative bodies of insurers and the shut down of their joint offices of statistics, one important advantage of large-scale business gets lost Not surprisingly, there has been a number of recent instances of failures of insurers. In fact, far more than the number of eventual wind-ups since many of them were hushed up by mergers

In these circumstances the solvency issue faces the actuarial profession with a number of challenging tasks. The appearance of actuaries of the third kind is a response to the problems associated with assets risk. In a sense these problems are harder than those associated with insurance liabilities assets risk is rooted in political, social, and economic phenomena of great complexity, whereas the fluctuations of insurance liabilities to a greater extent are governed by technical, physical, and demographic mechanisms that lend themselves to the well established methodology of the "exact sciences" This does not mean that the analysis of the liabilities is of secondary importance Just look at the classical life insurance mathematics Through decades it was widely held to be a largely perfect structure. However, it was not the mathematics that was perfect, but rather the idyll of the insurance companies in a situation where uniform premiums with substantial safety loadings built into them created great surplus. The insurers were prosperous and praised their actuaries The actuaries were flattered and praised their techniques. No development of theory was called for. Lately also life insurers are forced to compete, and suddenly the imperfection of the classical techniques is brought to light in confused discussions of how to determine appropriate premiums in different risk classes and how to redistribute surplus to them, in short, how to measure the risk Fresh thinking is required from all kinds of actuaries, first, second, and third, in order to meet the need for more accurate assessment of all kinds of risk in insurance In the present situation the only superfluous actuaries are those of the zero kind, who claim that actuarial mathematics can be dispensed with in these urgent matters.

It may be appropriate to coin the term "actuaries of the fourth kind" for those working in supervisory offices They are not numerous, and most of them lead a shadow life pondering returns from the company accountants Certainly, some very impressive work has been done in the field, but this fact alone could hardly justify a distinguishing mark It is the characteristics of the field itself and its great potential for stimulation of actuarial research that merits emphasis I shall list some items that hopefully will speak for themselves

— The objectives of the supervisory authorities are not all the same as those of a company. Solvency and equity are the primary concerns Business goals are balanced against the welfare of the insured, the efficiency of the insurance industry as a whole is considered, and its operations and organization can be influenced by statutory regulations Regardless of the market situation and the level of theoretical justification of the practices of actuaries of the three first kinds, the actuary of the fourth kind must employ models and methods that can serve these objectives (recall the life insurance situation) And when adequate theory does not exist, it must be created

— The data available to a supervisory office are different from those collected by the insurers Typically they are more aggregate and call for development of models at macro level and statistical methods based on these However, in our era of efficient data processing it is clearly possible to gather detailed statistics on policies and claims experiences for supervision purposes If this cannot be done on a large scale, an interesting possibility would be to study detailed data in carefully selected small samples from the insurance portfolios Then one can model at micro level, and derive the needed distributions for the totals determining the solvency state.

— The combination of data from several companies would presumably require employment of heterogeneity models to account for the unobservable differences between them The same goes also for the description of random fluctuation in collective risk factors Combining the two sources of variation leads to studies of two-way random effect models, not necessarily the standard linear ones.
— An important and difficult problem is the analysis of the impact of the size of the portfolio, its composition, and the reinsurance programme, which may be involved
— Yet another prominent problem is the projection of outstanding claims of all categories

The list of challenging actuarial and statistical problems could be extended far beyond this Some clues to their solutions are key-words like stochastic processes, prediction and filtering, finite time ruin probabilities in complex models, non- or semiparametric models, optimal risk sharing, utility and welfare theory, computerintensive statistical methods, standardization of definitions, organisation of statistical data bases and communication between these, ... Let it suffice here to say that all lines of insurance business have to be analysed statistically, and all aspects that are judged to be of significance to the total risk must be moulded into the analysis Not separately in ad hoc models, but simultaneously in one grand, comprehensive model, that must be sufficiently realistic and mathematically tractable to produce, on a large scale, reliable and efficient decisions in matters of major economic and social importance That is a formidable task and a great challenge to the actuarial science and profession.

RAGNAR NORBERG

XXI ASTIN COLLOQUIUM
NEW YORK, NOVEMBER 14-18, 1989

The XXI ASTIN Colloquium was held in the New York Hilton Hotel in central Manhattan The colloquium was attended by about 235 participants and 85 accompanying persons, coming from 22 countries. Approximately 25% of the participants were from the U.S.A., which is extraordinary but not surprising when one considers the location of the colloquium. The number of papers presented and contributions to the Speakers' Corner totalled 43.

The colloquium started informally with registration and a welcome drink in the evening of November 14. The official opening of the colloquium on November 15 coincided with the closing session of the 75th Jubilee Meeting of the Casualty Actuarial Society (CAS), and was held in the famous Waldorf Astoria Hotel. Kevin M. Ryan, President of CAS, held the opening address, followed by Jean Lemaire, Chairman of ASTIN Jean Lemaire handed over a congratulatory gift from ASTIN to Kevin M. Ryan on the occasion of the Casualty Actuarial Society's 75th jubilee. The Academy of Actuaries and the Conference of Actuaries in Public Practice, both actuarial associations in the U.S.A., had conveyed their welcome greetings to ASTIN, which were then read by Jean Lemaire

The first panel discussion of the session was on the *Past, Current and Future Role of Non-Life Actuaries Around the World.* Being videoed simultaneously, the discussion gave the audience an excellent opportunity to observe some eminent personalities in the actuarial profession at close range: LeRoy Simon the moderator, and Sidney Benjamin, Hans Bühlmann, Charles C. Hewitt and Jean Lemaire in the panel. Bühlmann gave a lucid account of the evolution of the actuarial species, culminating in the actuary of the third kind. Quite another kind of evolution was invoked by Hewitt, who described the impact of individual variations in mortality on population mortality rates, thus giving a plausible explanation of observed phenomena. Lemaire focused his speech on future challenges and opportunities of the actuarial profession; developments within the European Community (notably the planned recognition of actuarial qualifications across national boundaries); and the general decline of interest in mathematical studies, the traditional source of new recruits to the profession; and last but not least, the need to equip new actuaries with basic business skills

Benjamin sketched what one may call a code of conduct for the professional actuary with a special view to general insurance. The four panelists succeeded in giving quite a comprehensive view of actuarial preoccupations: Historical, strategic, technical and ethical issues were touched.

After the panel discussion the CAS meeting was ended officially with Kevin M. Ryan handing over the Presidency to Michael Fusco Members of CAS and ASTIN now reconvened to concurrent panel discussions with the following topics

1 Insurance Pricing: Return on Equity vs. Return on Sales. Yehuda Kahane, Bernard Pelletier, Richard Woll and David Hartman (moderator)
2. Pricing Tort Reform Robert Buchanan, Claus Metzner, Philip Miller and Paul Liscord (moderator).
3 Practical Applications of Determining Loss Developement Factors for Casualty Excess-of-Loss Business. Harold Clarke, Dan Lyons, Ben Zehnwirth and James MacGinnitie (moderator)

A delicious luncheon was then served at the New York Hilton Hotel, where the rest of the ASTIN Colloquium also was to take place Kevin M. Ryan addressed the participants at lunch, telling about the Casualty Actuarial Society's work in educational matters

After lunch the traditional working sessions started The referee would hardly be doing justice to the contributions by giving a one-sentence abstract of each paper. Thus I shall restrict myself to naming the contributants and the moderators of each session in chronological order, departing from this rule only when additional information seems interesting. Likewise I shall only name the presenting author of multi-author papers A complete list of papers and other contributions is given at the end

The first working session was moderated by Harri Lonka and Lionel Moreau. Papers were presented by Bob Alting von Geusau, Bob Buchanan, John Cozzolino, Chris Daykin, Bill Jewell and Stuart Klugman. Several papers presented in this session were results of cooperative effort. Buchanan's paper is a follow-up of a paper presented by Neuhaus at the 1985 ASTIN Colloquium in Biarritz Cozzolino, Klugman and Meyers presented their respective parts of work for the Insurance Services Office (ISO). The work presented by Daykin has connections to similar work done in Finland. Jewell's paper is part II in a trilogy on IBNYR reserving, part I of which was presented to the 1987 ASTIN Colloquium in Scheveningen.

The participants and their companions spent the evening on Broadway, or the Imperial Theater to be precise. The play of the evening was "Jerome Robbins' Broadway", a cavalcade of songs from musicals which Robbins staged over a period of 20 years. Listening to old favorites and watching a performance full of American precision, zest and humour made the evening thoroughly enjoyable

The following morning brought us back to the realm of actuarial mathematics Edwin J Elton, Nomura Professor of Finance at the Graduate School of Business, New York University, gave an invited survey lecture on the mathematical theory of investment The lecture was held in a very clear and concise style, giving the audience a glimpse of a vast new area for actuarial work, practical as well as theoretical

Thursday morning's working session was moderated by Marc Goovaerts. Jean Lemaire, Glenn Meyers and Ragnar Norberg presented their papers. Lemaire's paper, challenging actuaries to acquaint themselves with Fuzzy Set Theory, set off a lively debate involving Zehnwirth, Jewell, Norberg, Bühlmann and Hachemeister.

The ASTIN General Assembly followed. The minutes of the 1988 General Assembly, the Editor's report and the Treasurer's report were approved Robert Baumann of the Swiss Association of Actuaries announced the 1990 ASTIN Colloquium in Montreux, and Alf Guldberg of the Swedish Society of Actuaries announced the 1991 ASTIN Colloquium in Stockholm. A lengthy debate was generated by two suggested amendments to the ASTIN rules; both amendments were rejected with a clear majority vote. In the statutory elections Giovanna Ferrara and Ragnar Norberg withdrew from the ASTIN Committee and were replaced by Greg Taylor and Eddy Levay.

The first of the afternoon's working sessions was moderated by Richard Gauthier and Lars Austin Teivo Pentikainen, Bjørn Sundt, Hans Gerber, Maria de Lourdes Centeno and Erhard Kremer presented their work. Pentikäinen described the approach taken in the report *Insurance Solvency and Financial Strength* (abbr *The Blue Book*), a monumental work on insurance solvency done in Finland The Finnish solvency group has cooperated with the "Solvency Working Party" of the Institute of Actuaries (cf. Daykin's paper)

Charles Levi and Gary Patrik moderated the next afternoon session during which Lawrence Vitale, John Narvell & Peter Licht, Greg Taylor, Gary Venter and Walther Neuhaus presented their papers.

Friday was devoted entirely to working sessions. Eddy Levay moderated the first session with contributions from Gary Patrik, Steven Haberman, Yehuda Kahane, Charles Levi & Christian Partrat and Ermanno Pitacco.

Peter Johnson and Richard Gauthier moderated the second morning session with contributions from Mette Rytgaard, René Schnieper, Dirk Stiers, Bob van der Laan, and Alfred Weller.

The Friday morning sessions were strongly dominated by loss reserving, this being the subject of Kahane, Schnieper, Stiers and Weller. The lively discussion which arose between the advocates of elaborate models and the advocates of simple methods was all too often curtailed for lack of time The referee feels that the subject would have deserved a more thorough discussion which could have resolved at least superficial misunderstandings

Maurice R Greenberg, Chairman of the American International Group was the guest speaker during the luncheon. Greenberg conveyed an executive view on issues like environmental liability, proposing an alternative to current practice

Friday afternoon was devoted to the Speakers' Corner. Bill Jewell and Jean Casanova moderated the first session where papers were presented by João Manuel Andrade e Silva, Heikki Bonsdorff, Marc Goovaerts, Erhard Kremer and David Skurnick. Bonsdorff actually had two papers analyzing experience rating by exponential smoothing

John Narvell and Charles Hachemeister moderated the XXI ASTIN Colloquium's last working session Presentations were made by Ernesto Volpe, Robert Miccolis and Eugenio Prieto Perez. On a lighter note, Gunnar Benktander made a trendsetting remark, and Sidney Benjamin uttered "only one sentence" (sic) Martti Pesonen and Heikki Bonsdorff made further comments on *The Blue Book*; Bjørn Sundt proposed to systematise loss reserving acronyms. Teivo Pentikäinen, one of ASTIN's grand old men, struck the proper closing note by his stubborn pledge to carry on working May all our ASTIN colleagues be so stubborn!

The Colloquium Dinner took place on the 106th floor of the World Trade Center, with a spectacular view of New York and its suburbs. Jean Lemaire held the closing speech, thanking the American organisers for arranging the Colloquium in a superb way. Robert Baumann invited all ASTIN members to attend the XXII ASTIN Colloquium in Montreux. After dinner those who wished could dance to the orchestra. The largest ASTIN Colloquium to date, and one of the most interesting ones, had come to an end

For Saturday a tour of Manhattan, the United Nations and the Statue of Liberty had been arranged.

WALTHER NEUHAUS

LIST OF PAPERS AND OTHER CONTRIBUTIONS TO THE XXI ASTIN COLLOQUIUM

Risk Theory

Alting Von Geusau, B.: The Application of Additive and Multiplicative General Linear Interactive Models (GLIM) in Health Insurance

Buchanan, R.A., Heppell, I., Neuhaus, W.A.: A Hierarchical Credibility Model

Cozzolino, J.M.: Consistency of Risk Loaded Premiums

Daykin, C.D., Hey, G.B.: A Practical Risk Theory Model for the Management of Uncertainty in a General Insurance Company

Jewell, W.S.: Predicting IBNYR Events and Delays. II. Discrete Time

Klugman, S.A.: Measuring Uncertainty in Increased Limits Factors — A Bayesian Approach

Lemaire, J.: Fuzzy Insurance

Meyers, G.G.: A Supply Side Premium Calculation Principle

Norberg, R.: Outline of a Strategy for Solvency Control of Insurers

Pentikainen, T., Bonsdorff, H., Pesonen, M., Rantala, J., Ruohonen, M.: Evaluation of the Financial Strength of Insurers, Examples of Risk Theoretical Application

Willmot, G.E., Sundt, B.R.: On Evaluation of the Delaporte Distribution and Related Distributions

Gerber, H.U.: More Fun without Ruin Theory

Reinsurance

Centeno, M. de L.C., Simões, O.A.: Combining Quota-Share and Excess of Loss Treaties on the Reinsurance of *n* Independent Risks

Kremer, E.: The Asymptotic Efficiency of Largest Claims Reinsurance Treaties

Narvell, J.C., Licht, P.: Actuarial Involvement in the Liquidation of Insolvent Reinsurers

Taylor, G.C.: Optimal Reinsurance Structures

Venter, G.: Premium Calculation Implications of Reinsurance without Arbitrage

Vitale, L.A.: A Paper on the RAA Development Study

Neuhaus, W.A.: Mutual Reinsurance and Homogeneous Linear Estimation

Patrik, G., Mashitz, I.: Credibility for Reinsurance Excess Pricing

Empirical Studies

Haberman, S., Renshaw, A E : Fitting of Loss Distributions using Generalised Linear Models

Kahane, Y . A Modern Approach to Loss Reserving in Long-Tail Lines. The Case of Automobile Insurance.

Levi, C., Partrat, C.· Analyse statistique de catastrophes naturelles aux Etats-Unis

Pitacco, E. Adjustment Problems in Permanent Health Insurance

Rytgaard, M.· Estimation in the Pareto Distribution

Schmieper, R.: A Pragmatic IBNR Method

Stiers, D.: Applications de méthodes d'évaluation des réserves

Van der Laan, B.S., Hop, J P.· Probability Distributions for the Amount of Damage

Weller, A O.· Generalized Bondy Development

Silva, J.M A e: An Application of Generalized Linear Models to Portuguese Motor Insurance

Speakers' Corner

Bonsdorff, H.· On Changing the Parameter of Exponential Smoothing in Experience Rating

Bonsdorff, H : A Comparison of the Ordinary and a Varying Parameter Exponential Smoothing

Goovaerts, M An Algorithm for a Multi-Level Hierarchical Credibility Model

Kremer, E.: On the Probable Maximum Loss

Skurnick, D : Price Monitoring for Liability Insurance

Volpe, E.· Large and Catastrophic Risks

Benktander, G.. A Trendsetting Remark

Benjamin, S Only One Sentence!

Miccolis, R · Premium Principles based on Supply and Demand for Capital

Prieto Perez, E.. Preventive Measures An Analysis from the Point of View of their Efficiency

Pesonen, M.: Comments on the "Blue Book"

Bonsdorff, H More Comments on the "Blue Book"

Sundt, B R Comments on Loss Reserving Terminology

ARTICLES

THE ASYMPTOTIC EFFICIENCY OF LARGEST CLAIMS REINSURANCE TREATIES

By Erhard Kremer

Hamburg & Lohnberg, FRG

Abstract

Reinsurance treaties defined as generalizations of the classical largest claims reinsurance covers are investigated with respect to the associated risk, defined as the variance of the insurer's retaining total claims amount Instead of the unhandy variance corresponding handier asymptotic expressions are used With these an asymptotic efficiency measure for comparing two such reinsurance covers is defined. It is shown that with respect to asymptotic efficiency the excess-of-loss treaty is better than the classical largest claims treaty Furthermore the problem of giving optimal wheights to the ordered claims of a generalized largest claims cover is discussed.

INTRODUCTION

The choice of the appropriate treaty is a very old and fundamental problem in the reinsurance practice and theory Already in the sixties actuaries discussed the problem of the optimal choice of a reinsurance treaty. The stop-loss and quota shares were shown to have some very interesting optimality properties (see e.g Borch (1960), Kahn (1961), Lemaire (1973), Ohlin (1969), Pesonen (1967), Vajda (1962), Verbeek (1966) and the recent paper of Pesonen (1984)) Collective and individual treaties were compared and also an optimality property was given for the excess-of-loss treaty with respect to the class of individual treaties (see e.g Ohlin (1969), Gerber (1980)). A short presentation of these results is given e g in Kremer (1986b)

Nearly nothing is known on the goodness of the largest claims reinsurance treaty or of some of its interesting generalizations, which are defined e.g in Kremer (1986a), (1988a). Some remarks on certain dependencies between the largest claims and excess-of-loss treaty can be found in Berliner (1972). Furthermore one knows that under certain conditions the net premium of the classical largest claims cover (see e g Ammeter (1964)) is asymptotically equivalent to the net premium of a corresponding excess-of-loss treaty plus additive term (see Kremer (1982)) A generalization of that result to the generalized largest claims reinsurance covers was given by the author (see Kremer (1984)) some years ago Some more advanced results remain to show In the following the author presents some first new investigations on the goodness of the (classical or generalized) largest claims reinsurance treaties that are of the type one expects to get. Like in the already classical studies on the stop-loss, quota and excess-of-loss shares (see BORCH (1960), LEMAIRE (1973) and OHLIN (1969)) the author takes the inverse of the variance of the corresponding claims amount as measure for the goodness of the reinsurance treaty. Unfortunately one cannot give handy formulas for the variances under consideration. That's why the author replaces the variances by asymptotic formulas which were already cited in KREMER (1983). With these the asymptotic efficiency of two reinsurance treaties of the discussed type will be defined as the ratio of the inverses of the suitably transformed asymptotic variances. Like in the classical studies one takes the constraint that the net premiums of both treaties are (asymptotically) the same. With the help of this new concept of efficiency the classical largest claims cover (see AMMETER (1964)) is compared with the excess-of-loss treaty Finally the author deals with the problem of choosing optimally coefficients, weighting the ordered claims in the generalized largest claims reinsurance treaty.

THE GENERAL TREATY

Consider a collective of insurance risks producing claims with sizes $X_1, X_2, X_3, \ldots$, each year Denote with N the random variable describing the number of claims per year. The claims sizes are assumed to be stochastically independent and identically distributed with distribution function F. Finally the claims number is assumed to be independent of the claims sizes Investigated are reinsurance treaties which are based on the claims ordered in increasing amount i.e. on the random variables

$$X_{N1} \leqq X_{N2} \leqq \ldots \leqq X_{NN}$$

The reinsurance treaty conditions are defined by a family of weighting coefficients

$$c_{ni}, \qquad i = 1, 2, \ldots, n, \qquad n = 1, 2, 3 \ldots$$

and a function

$$h \text{ on the nonnegative reals}.$$

With these quantities the part of the total claims amount

$$S_N = \sum_{i=1}^{N} X_i$$

that the insurer retains, when concluding the treaty, is given according

$$R_N = \sum_{i=1}^{N} c_{Ni} \cdot h(X_{Ni}).$$

Such a reinsurance cover was recently called *linear reinsurance treaty based on ordered claims* (see e g KREMER (1988b)) In the more special situation where $h(x) = x$ holds for all x, one often denotes those reinsurance covers *generalized largest claims reinsurance treaties* (see e.g. KREMER 1988a)) The sense of the definition of the generalized treaty becomes obvious when considering some examples.

EXAMPLE 1. For the choice $c_{ni} = 1$ for all $i = 1, 2, 3, \ldots$ and $n = 1, 2, 3, \ldots$ and the special function

$$h(x) = \min(x, P)$$

with a given nonnegative priority P one gets the classical *excess-of-loss treaty* with priority P. The insurer has to pay for each claim up to the maximal amount P. ∇

EXAMPLE 2. In case that $c_{ni} = 1$ for all $i = 1, 2, \ldots, n-p$ and $c_{ni} = 0$ otherwise and that $h(x) = x$ holds for all x, one has the (classical) *largest claims reinsurance treaty*, where the reinsurer pays for all claims except the p largest ones ∇

The reader is invited to give some more examples, e.g one can combine the situations of the example 1 and 2. Notice that in the present investigations we consider the claims amount remaining by the insurer and not like in the previous studies (see KREMER (1986a), (1988a), (1988b)) the claims amount taken by the reinsurer. In other words, the R_N here is just the $S_N - R_N$ of the previous papers.

THE ASYMPTOTIC EFFICIENCY

Obviously the class of linear reinsurance treaties based on ordered claims is fairly large One can choose among many different such reinsurance covers The question appears which of two given different treaties is preferable. For deciding, one needs an appropriate measure with which one can select the treaty which is more advantageous. A classical measure for judging the goodness of a reinsurance treaty is the variance of the total claims amount under consideration while choosing some parameters of the treaty such that the mean value of the total claims amount is fixed (see e.g. OHLIN (1969) and KREMER (1986b) chapter 5.1). In case of our above defined linear reinsurance treaty based on ordered claims no handy expressions for the expectation $E(R_N)$ and variance $\mathrm{Var}(R_N)$ exist in general. Fortunately one can give elegant formulas for both quantities with asymptotic considerations. More concretely the author gave in a previous paper (see KREMER (1983)) expressions for

$E(R_N)$ and $\operatorname{Var}(R_N)$ that are asymptotically equivalent to both quantities These results are basic to all that follows and will be presented in the sequel.

Consider a sequence of growing collectives, indexed with the integer $k = 1, 2, 3, \ldots$ Denote with N_k the claims number of the collective no. k and suppose that

$$\lim_{k \to \infty} (E(N_k)) = +\infty$$

$$\lim_{k \to \infty} \left(\frac{\operatorname{Var}(N_k)}{E(N_k)} \right) = c$$

with an arbitrary, but fixed constant c. The random variables of the claims amounts are the same in each collective and denoted by the variables

$$X_1, X_2, X_3, \ldots$$

They are assumed to satisfy the conditions given in the beginning of the previous section, especially they are assumed to be stochastically independent of the claims numbers N_k and to have the distribution function F The linear reinsurance treaty based on ordered claims now depends also on the collective number k, more concretely the weighting coefficients are dependent of the index number k:

$$c_{ni} = c_{ni}^{(k)}, \qquad \text{with} \qquad k = 1, 2, 3, \ldots,$$

whereas the function h is independent of the number k. For giving the asymptotic formulas for the expectation and variance of the claims amount

$$R_N = R_{N_k}^{(k)} = \sum_{i=1}^{N_k} c_{N_k i}^{(k)} \, h(X_{N_k, i}),$$

one defines the family of functions

$$b_n^{(k)}, \qquad n = 1, 2, 3, \ldots, \qquad k = 1, 2, 3,$$

according

$$b_n^{(k)}(0) = c_{n1}^{(k)},$$

$$b_n^{(k)}(u) = c_{ni}^{(k)} \text{ for } u \text{ in the interval } ((i-1)/n, i/n] \quad \text{and} \quad i = 1, 2, 3, \ldots, n,$$

and assumes that there exist an *asymptotic weighting function* b and numbers

$$t_i, \qquad i = 0, 1, \ldots, m+1$$

with $t_i < t_{i+1}$, $t_0 = 0$, $t_{m+1} = 1$ such that

$$\lim_{k\to\infty} (b_{n_k}^{(k)}) = b$$

uniformly on closed subintervals of the complementary set of $\{t_1, \ldots, t_m\}$ and for each sequence $(n_k, k = 1, 2, \ldots)$ satisfying

$$\lim_{k\to\infty} \left(\frac{n_k}{E(N_k)}\right) = 1 .$$

The function b is supposed to consist of two parts

$$b = b_s + b_d ,$$

where b_s is of bounded variation and continuously differentiable and b_d is a step function with steps at the points $t_1, \ldots, t_m$. Finally the function h shall be nondecreasing, F be continuous and strictly increasing (from both sides) at the points $F^{-1}(t_i)$, $i = 1, 2, \ldots, m$, with the convention

$$F^{-1}(u) = \inf\{x : F(x) \geqq u\}$$

With all these notations and assumptions one has the important result that with the expressions

(1) $$\mu_F(b, h) = \int_0^\infty b(F(x)) \cdot h(x)\, F(dx)$$

(2) $$\sigma_F^2(b, h) = \int_0^\infty \int_0^\infty (\min(F(s), F(t)) - F(s) \cdot F(t)) \times \times b(F(s)) \cdot b(F(t))\, h(\mathrm{d}s)\, h(dt)$$

holds

(3) $$\lim_{k\to\infty} \left(\frac{E(R_{N_k}^{(k)})}{E(N_k)}\right) = \mu_F(b, h)$$

(4) $$\lim_{k\to\infty} \left(\frac{\operatorname{Var}(R_{N_k}^{(k)})}{E(N_k)}\right) = \sigma_F^2(b, h) + c \cdot \mu_F^2(b, h)$$

(see Theorem 1 in KREMER (1983)). In these formulas the distribution function F is fixed and given. The expressions depend only through the functions b and h on the linear reinsurance treaty based on ordered claims.

Now coming back to judging the goodness of a linear reinsurance treaty based on ordered claims. In the classical approach of comparing reinsurance treaties one fixes the expectation of the claims amount under consideration and investigates the corresponding variance. According to the elegant result (3) the fixing of the expectation can be formulated in an asymptotic sense according:

$$\mu_F(b, h) = \text{constant}.$$

Then according to (4) the investigation of the variance can be expressed in an asymptotic sense as the investigation of the expression

$$\sigma_F^2(b, h).$$

All above remarks now are summarized in the following definition.

DEFINITION. In the above setting consider two linear reinsurance treaties based on ordered claims with corresponding functions b_i, h_i, $i = 1, 2$ Suppose that

(5) $$\mu_F(b_1, h_1) = \mu_F(b_2, h_2)$$

is satisfied. Then the value

$$EFF_F(1:2) = \left(\frac{\sigma_F^2(b_2, h_2)}{\sigma_F^2(b_1', h_1)}\right)$$

is called *asymptotic efficiency* of the treaty no. 1 relative to the treaty no 2. In case that

(6) $$EFF_F(1 \cdot 2) > 1$$

holds true, the treaty no. 1 is called *better* than the treaty no 2 (at the underlying claims size distribution function F). In case that in (6) one has equality, both treaties are called to be (asymptotically) *equivalent*. ∇

Obviously this definition gives a practicable formal instrument for comparing the linear reinsurance treaties based on ordered claims. In case of the socalled generalized largest claims treaties (see above) one has $h(x) = x$ for all x, so that μ_F, σ_F^2 depend only on b. Then write shorter $\mu_F(b)$, $\sigma_F^2(b)$ for the special $\mu_F(b, h)$, $\sigma_F^2(b, h)$. For illustration an important example shall be discussed.

AN EXAMPLE

Consider the (classical) largest claims reinsurance treaty of the example 2 in the above context of growing collectives. Denote the number p of the treaty in the collective no k by p_k. Assume that

$$\lim_{k \to \infty} \left(\frac{p_k}{E(N_k)}\right) = s,$$

for an arbitrary, fixed value s between zero and one. This treaty shall be compared with the excess-of-loss treaty with priority $P > 0$ (see the example 1). With these notations one gets the corresponding functions.

(a) *for the excess-of-loss treaty*

$$b_1(u) = 1, \qquad \text{for all } u,$$

$$h_1(x) = \min(x, P), \quad \text{for all } x,$$

(b) *for the largest claims treaty*

$$\begin{aligned} b_2(u) &= 1, \text{ if } u \text{ is smaller or equal to } 1-s, \\ &= 0, \text{ elsewhere,} \end{aligned}$$

$$h_2(x) = x, \text{ for all } x,$$

and the interesting result:

THEOREM The excess-of loss treaty is better than the largest claims cover in the just given setting ∇

PROOF. Since

$$\mu_F(b_1, h_1) = \int_{[0,\, P]} x\, F(dx) + P\ (1-F(P))$$

$$\mu_F(b_2, h_2) = \int_{[0,\, F^{-1}(1-s)]} x\, F(dx)\,,$$

the equation (5) means nothing else but that

$$\int_{[0,\, P']} x\, F(dx) - \int_{[0,\, P]} x\, F(dx) = P\cdot(1-F(P)) \tag{7}$$

with the priority $P' = F^{-1}(1-s)$ This implies at once that $P' \geqq P$ Since $F(P') < 1$, one also has $F(P) < 1$. Consequently one has because of (7) the stronger condition:

$$P' > P \tag{8}$$

Inserting the b_i, h_i into the expression (2) implies the in structure identical formulas:

(9) $$\sigma_F^2(b_1, h_1) = 2 \int_0^P F(r) \cdot \int_r^P (1 - F(t))\, dt\, dr$$

(10) $$\sigma_F^2(b_2, h_2) = 2 \cdot \int_0^{P'} F(r) \int_r^{P'} (1 - F(t))\, dt\, dr\,,$$

from which one concludes easily with (8) that the excess-of-loss treaty is the better one. ∇

This result can be seen as a theoretical justification for the common preference of the excess-of-loss treaty on the international reinsurance market. Surprisingly the proof of the theorem is fairly simple, when using the concepts of the preceeding section. The result is new and fits well to the investigations of BERLINER (1972). For given $s \in (0, 1)$ and distribution function F one can easily compute the efficiency of the largest claims treaty relative to the excess-of-loss treaty. Since s is given, the P' is fixed, so that one can compute the corresponding priority P from the equation (7). For computing $EFF_F(1:2)$ it then remains to evaluate the integrals in (9) and (10) and then to take the ratio Exemplarily one can take for F the classical Pareto-model

COROLLARY Suppose that F is the Pareto-model, i.e

$$F(x) = 1 - x^{-a}, \text{ for } x \text{ larger than } 1\,,$$

with a given parameter a larger than 2. Define the following function:

$$g(y) = y^{2(1-a)} \left(\frac{1}{2 \cdot (a-1)}\right) + y^{2-a} \left(\frac{a-1}{a-2}\right) - $$

$$- y^{1-a} \cdot \left(\frac{a}{a-1}\right) - \frac{a}{2 \cdot (a-1) \cdot (a-2)}$$

and the values:

$$y1 = a^{1/(1-a)}\, s^{-1/a}$$
$$y2 = s^{-1/a}$$

With this notation one gets for the asymptotic efficiency of the largest claims relative to the excess-of-loss treaty the result

$$EFF_F(1:2) = \frac{g(y1)}{g(y2)}\,,$$

in case that s is smaller than $a^{a/(1-a)}$.

PROOF One evaluates with routine calculations the equations (7), (9) and (10). More concretely (7) means with $y1 = P$, $y2 = P'$ that:

$$a \cdot \int_{y1}^{y2} x^{-a}\, dx = y1^{1-a}$$

what is equivalent with

$$y1 = a^{1/(1-a)} \cdot y2 .$$

Since:

$$P' = F^{-1}(1-s) = s^{-1/a},$$

one has the formulas for $y1 = P$ and $y2 = P'$. Furthermore one shows that.

$$\int_0^y F(r) \cdot \int_r^y (1-F(t))\, dt\, dr$$

$$= \int_1^y (1-r^{-a}) \cdot \left(\int_r^y t^{-a}\, dt \right) dr = \ldots = g(y) \cdot (1-a)^{-1} \cdot \nabla$$

The Corollary shows that the efficiency depends for given parameters of the distribution function F solely on the value of s.

OPTIMAL WEIGHTS

In this section the problem of how to choose the *weighting coefficients* c_{ni}, $i = 1, 2, \ldots, n$, $n = 1, 2, \ldots$ is discussed for the situation that the function h of the linear reinsurance treaty based on ordered claims is given and the insurer likes to retain a net premium exceeding a minimum amount μ Without loss of generality let us assume that

$$h(x) = x, \qquad \text{for all } x .$$

This means that one deals with the generalized largest claims reinsurance treaty and tries to find some in some sense optimal weighting coefficients for given claims size distribution function F and on a constraint on the insurer's net income. Suppose that one has one in some sense optimal *asymptotic weighting function* b for the treaty. Then an adequate choice of the weighting coefficients is to take

$$c_{ni} = b\left(\frac{i}{n}\right), \qquad \text{for} \qquad i = 1, 2, \ldots, n$$

$$n = 1, 2, \ldots$$

With this choice the treaty is in some sense asymptotically optimal So the problem of giving adequate weighting coefficients reduces to the problem of determining an optimal asymptotic weighting function. The above presented concepts and ideas make it possible to define what might be regarded as an optimal asymptotic weighting function b

DEFINITION. Consider the class of generalized largest claims reinsurance treaties with asymptotic weighting functions b in the above context of growing collectives. Suppose that one has with a given constant μ the restriction on b:

$$\mu_F(b) \geqq \mu\,, \tag{11}$$

where F is the fixed underlying distribution function of the claims sizes of the collective. The asymptotic weighting function b_* is called *optimal* in the class Δ if·

$$\bar{\sigma}_F^2(b_*) = \inf_{b \in \Delta} [\bar{\sigma}_F^2(b)]\,,$$

where Δ is a given class of asymptotic weighting functions with each $b \in \Delta$ satisfying (11) and the $\bar{\sigma}_F^2(b)$ is the right hand side of (4). A treaty with corresponding asymptotic weighting function b_* is called *asymptotically optimal* in the class of treaties with weighting functions $b \in \Delta$. ∇

Assuming that the basic claims size distribution function F is continuous, one can reformulate $\mu_F(b)$ and $\sigma_F^2(b)$ according

$$\mu_F(b) = \int_0^1 b(u) \cdot F^{-1}(u)\, du$$

$$\sigma_F^2(b) = \int \int (\min(u, v) - u\ v) \cdot b(u)\ b(v)\, F^{-1}(du)\, F^{-1}(dv)\,.$$

Obviously σ_F^2 is something like a quadratic form and the μ_F nothing else but a linear functional on the set of all asymptotic weighting functions b. In case one restricts to uniformly continuous functions b, the determination of an optimal b, reduces to a typical infinite optimization problem, i e to the minimization of the sum of a quadratic and a squared linear form under the contraint that a linear functional exceeds a given constant For solving one can apply results of the socalled infinite optimization theory. The reader is referred to the literature on this mathematical topic (see e.g KRABS (1975)) Because of practical reasons one will take in addition the condition on b that

$b(u)$ is nonnegative for all u and bounded by the amount 1.

One knows that the optimal b_* can be determined such that with a nonnegative λ_* the tupel (b_*, λ_*) is a saddlepoint of the Lagrange-function

$$L(b, \lambda) = \bar{\sigma}_F^2(b) - \lambda \cdot (\mu_F(b) - \mu)$$

with respect to all nonnegative b and λ

In practice one does not know F but only knows the corresponding empirical distribution function F_m, defined with the known m past claims amounts $X_1, X_2, \ldots, X_m$ according·

$$F_m(x) = \left(\frac{1}{m}\right) \cdot (\text{number of } X_i \leqq x) .$$

Then one clearly inserts F_m for F in the $\mu_F(b)$, $\sigma_F^2(b)$, yielding as results

$$(12) \qquad \mu_{F_m}(b) = \left(\frac{1}{m}\right) \sum_{i=1}^{m} b\left(\frac{i}{m}\right) \cdot X_{m\,i}$$

$$(13) \qquad \sigma_{F_m}^2(b) = \left(\frac{1}{m^2}\right) \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} (m \min(i,j) - i \cdot j)) \times$$

$$\times\, b\left(\frac{i}{m}\right) b\left(\frac{j}{m}\right) (X_{m\,(i+1)} - X_{m\,i}) \cdot (X_{m\,(j+1)} - X_{m\,j})$$

with probability one. Here one uses that for continuous F the ordered claims are all different, i e

$$(14) \qquad X_{m\,1} < X_{m\,2} < . \quad < X_{m\,m},$$

with probability one. In case some ordered claims are equal, both expressions in (12), (13) have to be modified slightly. Let us restrict exemplarily to the situation (14) In (12), (13) the asymptotic weighting function $b(u)$ appears only at the points $u = i/m$, where i runs from 1 to m One can calculate optimal values $b_1, b_2, . ., b_m$ for $b(1/m), b(2/m), \quad, b(1)$ by minimizing $\bar{\sigma}_{F_m}^2(b)$ with respect to the $b(i/m)$ (with $i = 1, 2, .. \ , m$) under the constraints that

$$\mu_{F_m}(b) \geqq \mu .$$

and that the $b(i/m)$ are nonnegative and bounded by 1 This is nothing but a standard problem of the (finite) optimization theory, which can be solved with the methods of the *quadratic programming* (see e g. KUNZI et al. (1967)).

Having calculated the optimal values $b_*(i/m) = b_i$ with $i = 1, 2, . \ , m$, one likes to have also values $b_*(u)$ for the u unequal to the i/m with $i = 1, 2, \ldots, m$ A practical approach might the simply to interpolate and extrapolate the function $b_*(u)$ between and from the points $u = i/m$ with $i = 1, 2, \ . \ , m$, by using a suitable method of the numerical mathematics.

Since the methods of the quadratic programming nowadays work without great problems on each modern computer, one can determine with the given procedure an approximate optimal asymptotic weighting function b_* in case one has the empirical distribution function.

Clearly the problem of giving optimal weights c_{ni} $(i = 1, 2, \ldots, n,$ $n = 1, 2, \ldots)$ or more concretely an optimal asymptotic weighting function b is mainly of theoretical interest In practice the reinsurer clearly will not loose time with computing such an "optimal" reinsurance treaty So the author restricts to the above short discussion and closes the paper

REFERENCES

AMMETER, H (1964) The rating of "largest claim" reinsurance covers *Quaterly letter from the allgemeene reinsurance companies*, 79–109
BEARD, R, PENTIKANEN, P and PESONEN, E (1969) *Risk Theory* Chapman & Hall, Londen
BERLINER, B (1972) Correlations between excess-of-loss reinsurance covers and reinsurance of the n largest claims *ASTIN Bulletin*, 260 275
BORCH, K (1960), An attempt to determine the optimum amount of stop loss reinsurance *Transactions of the XVI international congress of actuaries*, 597 610
GERBER, H U (1980) *An introduction to mathematical risk theory* S S Huebner foundation, Philadelphia
KAHN, P (1961) Some remarks on a recent paper by Borch *ASTIN Bulletin*, 265–272
KRABS, W (1975) Optimierung und Approximation *Teubner Studientexte*, Stuttgart.
KREMER, E (1982), Rating of largest claims and ECOMOR reinsurance treaties for large portfolios *ASTIN Bulletin*, 47–56
KREMER, E (1983), Rating of nonproportional reinsurance treaties based on ordered claims A lecture given at the NATO-study institute in Leuven Published in 1984 in the *NATO-ASI series Premium Calculation in Insurance*, edited by Goovaerts, M, De Vylder, F and Haezendonck, J
KREMER, E (1984) An asymptotic formula for the net premium of some reinsurance treaties *Scandinavian Actuarial Journal*, 11–22
KREMER, E (1986a) Recursive calculation of the net premium of largest claims reinsurance covers *ASTIN Bulletin*, 101–108
KREMER, E (1986b) *Hohere Versicherungsmathematik* Verein zur Forderung der angewandten mathematischen Statistik und Risikotheorie Hamburg
KREMER, E (1988a) A general bound for the net premium of largest claims reinsurance treaties *ASTIN Bulletin*, 69-78
KREMER, E (1988b) *Net premiums of nonproportional reinsurance covers in nonlife insurance* Verein zur Forderung der angewandten mathematischen Statistik und Risikotheorie Hamburg
KÜNZI, H P, TZSCHACH, H G and ZEHNDER, C A (1967) *Mathematische Optimierung* Teubner Studienbucher, Stuttgart
LEMAIRE, J (1973) Sur la détermination d'un contract optimal de réassurance *ASTIN Bulletin*, 165–180
OHLIN, J (1969) On a class of measures of dispersion with application to optimal reinsurance *ASTIN Bulletin*, 249-266
PESONEN, E (1967) On optimal properties of the stop loss reinsurance *ASTIN Bulletin*, 175–176
PESONEN, M I (1984) Optimal reinsurances *Scandinavian Actuarial Journal*, 65–90
VAJDA, S (1962) Minimum variance reinsurance *ASTIN Bulletin*, 257–260
VERBEEK, H G (1966) On optimal reinsurance *ASTIN Bulletin*, 29 38

Prof. Dr. ERHARD KREMER
Verein zur Förderung der Angewandten Mathematischen Statistik und Risikotheorie
Research & Relax Area, Wallstr. 15, 6293 Löhnberg 1, FRG.

PARETO OPTIMAL RISK EXCHANGES AND A SYSTEM OF DIFFERENTIAL EQUATIONS: A DUALITY THEOREM

By Erich Wyler

ETH Zürich, Switzerland

Abstract

This article, based on a result of Borch and an extension of Bühlmann, gives a complete characterization of Pareto optimal risk exchanges by a system of differential equations linking the derivate of agents contributions to their risk aversion coefficients.

Keywords

Pareto optimal risk exchange; Bernoulli utility function, absolute risk aversion, system of differential equations.

1. Introduction

This article extends a result of Bühlmann (1984) Starting from Borch's theorem (1960), Bühlmann found a system of differential equations with a Pareto optimal risk exchange as the solution. Here we are starting from these differential equations and prove existence and uniqueness of a solution without assuming any further condition. This solution depends on initial values which satisfy a certain clearing condition. It will turn out that it can be identified in a bijective way with the set of Pareto optimal risk exchanges

2. Model

We consider a risk pool with n participants. Participant i $(1 \leq i \leq n)$ is characterized by

- r_i : initial wealth
- X_i : initial risk (random variable defined on a probability space $(\Omega, \mathfrak{A}, P)$; we assume that the expected values $E[X_i]$ exists)
- u_i : Bernoulli utility function (defined on $\mathbb{R}$, increasing, strictly concave and twice differentiable: $u_i' < 0$, $u_i'' > 0$)
- ρ_i : absolute risk aversion ($\rho_i := -u_i''/u_i'$. Notice $u_i' > 0$ and $\rho_i > 0$ i.e. the participants are risk averse; see Pratt (1964)).

By a risk pool we mean any formal mutual agreement among the n participants to redistribute their total initial risk $\sum_{i=1}^{n} X_i$.

The intial risk vector X,

$$X := (X_1, \ldots, X_n),$$

is called risk vector before exchange whereas a risk vector Y,

$$Y := (Y_1, \ldots, Y_n),$$

defined on the same probability space $(\Omega, \mathfrak{A}, P)$ and satisfying the clearing condition

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} X_i ,$$

is denoted as risk vector after exchange or briefly as risk exchange

Furthermore a risk exchange $Y^* := (Y_1^*, \ldots, Y_n^*)$ is called Pareto optimal if there does not exist another risk exchange $Y := (Y_1, \ldots, Y_n)$ with

$$E[u_i(r_i - Y_i^*)] \leq E[u_i(r_i - Y_i)] \qquad \text{for all } i$$

$$E[u_{i^o}(r_{i^o} - Y_{i^o}^*)] < E[u_{i^o}(r_{i^o} - Y_{i^o})] \qquad \text{for at least one } i^o.$$

In the sequel we are interested in Pareto optimal risk exchanges

REMARK The motivation of a person for participating in a risk pool is to improve his initial expected utility $E[u(r-X)]$ Therefore a risk exchange Y has to satisfy the individual rationality condition

$$E[u_i(r_i - X_i)] \leq E[u_i(r_i - Y_i)] \qquad \text{for all } i$$

in addition to the pool condition of Pareto optimality Unfortunately there are many Pareto optimal risk exchanges violating this condition. In order to preserve the beauty of the main result we drop the individual rationality condition and deal in this article with general Pareto optimal risk exchanges.

In order to simplify our notation we introduce the shifted disutility functions v_i

$$v_i(x) := u_i(r_i - x) \qquad i = 1, \ldots, n \qquad (v_i' < 0,\ v_i'' < 0)$$

With W_i we denote the range of the derivative of v_i

$$W_i := \{v_i'(x) \mid x \in \mathbb{R}\}$$

3. MAIN RESULT

Now we show the existence of a bijective mapping between the set of Pareto optimal risk exchanges and the set of solutions of a system of differential equations satisfying a constrained boundary condition.

THEOREM

Let w, $w := (w_1, \ldots, w_n) \in \mathbb{R}^n$, be a vector with $\sum_{i=1}^{n} w_i = 0$

(i) Let (A) be the system of differential equations

$$\text{(A)} \qquad Y_i'(z) = \frac{\dfrac{1}{\rho_i(r_i - Y_i(z))}}{\sum_{j=1}^{n} \dfrac{1}{\rho_j(r_j - Y_j(z))}} \qquad i = 1, \ldots, n$$

There exists a uniquely defined solution $Y(z) = (Y_1(z), \ldots, Y_n(z))$ of (A) satisfying the boundary condition $Y_i(0) = w_i$, $i = 1, \ldots, n$

(ii) If $Y(z) = (Y_1(z), \ldots, Y_n(z))$ is the solution of (A) with boundary condition $Y_i(0) = w_i$, $i = 1, \ldots, n$, then

$$Y\left(\sum_{i=1}^{n} X_i\right)$$

is a Pareto optimal risk exchange.

(iii) If $Y^* := (Y_1^*, \ldots, Y_n^*)$ is a Pareto optimal risk exchange then there exists a solution $Y(z) = (Y_1(z), \ldots, Y_n(z))$ of (A) satisfying a uniquely defined boundary condition $Y_i(0) = w_i$, $i = 1, \ldots, n$, $\sum_{i=1}^{n} w_i = 0$, with

$$Y^* = Y\left(\sum_{i=1}^{n} X_i\right) \qquad \text{almost surely.}$$

PROOF

(i) Existence of a solution:

Let f_k be the function $f_k(x) = \sum_{i=1}^{n} (v_i')^{-1}\left(\dfrac{x}{k_i}\right)$ with $k := (k_1, \ldots, k_n)$,

$k_i := \dfrac{-1}{v_i'(w_i)} > 0$. f_k is a strictly decreasing and differentiable function defined on W

$$W = \bigcap_{i=1}^{n} \{x\, k_i \mid x \in W_i\}$$

with range $\mathbb{R}$ (see Lemma 1, Appendix). Furthermore $f_k(-1) = 0$. (see Proof of Lemma 1, Appendix) We have

$$Y(z) = (Y_1(z), \ldots, Y_n(z)) \text{ with } Y_i(z) = (v_i')^{-1}\left(\frac{1}{k_i}(f_k)^{-1}(z)\right), \; i = 1, \ldots, n,$$

and $k_i := \dfrac{-1}{v_i'(w_i)}$, $i = 1, \ldots, n$, is a solution of (A)

Uniqueness of the solution:

Let $\tilde{Y}(z) = (\tilde{Y}_1(z), \ldots, \tilde{Y}_n(z))$ be another solution of (A) satisfying the same boundary condition. We define differentiable functions $g_i(z)$, $i = 2, \ldots, n$:

$$g_i(z) := k_1 v_1'(\tilde{Y}_1(z)) - k_i v_i'(\tilde{Y}_i(z)), \qquad z \in \mathbb{R}$$

We have $g_i(0) = 0$ for all i and for the derivatives $g_i'(z)$, $i = 2, \ldots, n$, we get

$$g_i'(z) = k_1 v_1''(\tilde{Y}_1(z))\, \tilde{Y}_1'(z) - k_i v_i''(\tilde{Y}_i(z))\, \tilde{Y}_i'(z)$$

$$= \frac{\dfrac{k_1 v_1''(\tilde{Y}_1(z))}{\rho_1(r_1 - \tilde{Y}_1(z))} - \dfrac{k_i v_i''(\tilde{Y}_i(z))}{\rho_i(r_i - \tilde{Y}_i(z))}}{\displaystyle\sum_{j=1}^{n} \frac{1}{\rho_j(r_j - \tilde{Y}_j(z))}} \quad \text{(with (A))}$$

$$= \frac{g_i(z)}{\displaystyle\sum_{j=1}^{n} \frac{1}{\rho_j(r_j - \tilde{Y}_j(z))}}, \quad z \in \mathbb{R}.$$

Because the homogeneous linear differential equations

$$g_i'(z) = \frac{g_i(z)}{\displaystyle\sum_{j=1}^{n} \frac{1}{\rho_j(r_j - \tilde{Y}_j(z))}}, \quad z \in \mathbb{R}, \quad i = 2, \ldots, n$$

have only solutions of the form

$$g_i(z) = c_i \exp\left(\int_0^z \frac{1}{\displaystyle\sum_{j=1}^{n} \frac{1}{\rho_j(r_j - \tilde{Y}_j(t))}}\, dt \right), \quad c_i \in \mathbb{R}, \quad i = 2, \ldots, n$$

we get together with $g_i(0) = 0$: $c_i = 0$ and therefore $g_i(z) = 0$ for all $z \in \mathbb{R}$ and $i = 2, \ldots, n$.

This means

$$k_i v_i'(\tilde{Y}_i(z)) = k_1 v_1'(\tilde{Y}_1(z)) \quad \text{for all} \quad z \in \mathbb{R} \quad \text{and} \quad i = 2, \ldots, n.$$

Because $\sum_{i=1}^{n} \tilde{Y}_i'(z) = 1$ for all $z \in \mathbb{R}$ it follows together with the boundary condition that $\sum_{i=1}^{n} \tilde{Y}_i(z) = z$ for all $z \in \mathbb{R}$

Because $\tilde{Y}(z)$ and $Y(z)$ satisfy both the equations (**) of Lemma 2 (see Appendix) we conclude by uniqueness of the solution that

$$\tilde{Y}(z) = Y(z) \qquad \text{for all} \qquad z \in \mathbb{R}\,.$$

(ii) $Y(z) := (Y_1(z), \ldots, Y_n(z))$ with $Y_i(z) = (v_i')^{-1}\left(\frac{1}{k_i}(f_k)^{-1}(z)\right)$, $i = 1, \ldots, n$,

and $k_i := \frac{-1}{v_i'(w_i)}$, $i = 1, \ldots, n$, is the unique solution of (A)

satisfying the boundary condition $Y_i(0) = w_i$, $i = 1, \ldots, n$. Because this solution satisfies (**) of Lemma 2, (see Appendix), it follows from BORCH'S theorem (see BORCH, (1960))that $Y(\Sigma X_i)$ is a Pareto optimal risk exchange

(iii) It follows from BORCH'S theorem (see BORCH, (1960)) that there are strictly positive constants k_i, $i = 1, \ldots, n$, with

$$k_i v_i'(Y_i^*) = k_1 v_1'(Y_1^*) \quad \text{almost surely for} \quad i = 2, \ldots, n\,.$$

Let $\omega \in \Omega$ be an element of Ω for which the condition of BORCH is satisfied, $\boldsymbol{k}$ the vector $\boldsymbol{k} := (k_1, \ldots, k_n)$ and f_k the function as defined above. Because $f_k(x)$ is defined for $x := k_1 v_1'(Y_1^*(\omega))$

$$f_k(x) = \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{x}{k_i}\right) = \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{k_i v_i'(Y_i^*(\omega))}{k_i}\right) = \sum_{i=1}^{n} Y_i^*(\omega)$$

(condition of BORCH)

it follows analogously to Lemma 1, (see Appendix), that f_k is defined on some interval (a, b) with range $\mathbb{R}$. Therefore $(f_k)^{-1}(0)$ exists. We define the vector $\boldsymbol{w} = (w_1, \ldots, w_n)$ by

$$w_i := (v_i')^{-1}\left(\frac{1}{k_i}(f_k)^{-1}(0)\right), \qquad i = 1, \ldots, n\,.$$

The unique solution $Y(z) = (Y_1(z), \ldots, Y_n(z))$ of (A) with boundary condition $Y_i(0) = w_i$, $i = 1, \ldots, n$, satisfies the equations (**) of Lemma 2, i.e.

$$\sum_{i=1}^{n} Y_i(z) = z \qquad \text{for all} \qquad z \in \mathbb{R}$$

$$k_i v_i'(Y_i(z)) = k_1 v_1'(Y_1(z)) \quad \text{for} \quad i = 2, \ldots, n \quad \text{and all} \quad z \in \mathbb{R}\,.$$

We conclude by uniqueness of the solution that

$$Y^* = Y\left(\sum_{i=1}^{n} X_i\right) \qquad \text{almost surely}\,.$$

(REMARK: Because $-1/k_i$ is possibly not in the range of v_i' we cannot

define w_i by $w_i := (v_i')^{-1}\left(\frac{-1}{k_i}\right)$) QED

4 EXAMPLE

We assume that the participants are using exponential utility functions, i.e. $\rho_i(x) = a_i$ for all $x \in \mathbb{R}$ and $i = 1, \ldots, n$, where ρ_i denotes the absolute risk aversion of participant i. In this case the system of differential equations (A) becomes very simple

$$\text{(A)} \qquad Y_i'(z) = \frac{\frac{1}{a_i}}{\sum_{j=1}^{n} \frac{1}{a_j}}, \qquad i = 1, \ldots, n$$

We therefore have

$$Y_i(z) = \frac{\frac{1}{a_i}}{\sum_{j=1}^{n} \frac{1}{a_j}} z + \beta_i, \qquad i = 1, \ldots, n,$$

where the β_i's satisfy the clearing condition

$$\sum_{i=1}^{n} \beta_i = 0 .$$

For further examples, e.g. for utility functions of the HARA-type, see LIENHARD, (1986).

APPENDIX

To conclude the two technical lemmas already used in the proof of the main Theorem are discussed.

LEMMA 1

Let $\boldsymbol{w}, \boldsymbol{w} := (w_1, \ldots, w_n) \in \mathbb{R}^n$, be a vector with $\sum_{i=1}^{n} w_i = 0$

and $\boldsymbol{k}, \boldsymbol{k} := (k_1, \ldots, k_n) \in \mathbb{R}^n$ the vector with $k_i = \frac{-1}{v_i'(w_i)} > 0$

Then $f_k(x) := \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{x}{k_i}\right)$ is a strictly decreasing and differentiable function defined on W

$$W := \bigcap_{i=1}^{n} \{x k_i \mid x \in W_i\}$$

with range $\mathbb{R}$.

PROOF

Obviously W is an open interval. W is not empty because it contains -1:

$$f_k(-1) = \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{-1}{k_i}\right) = \sum_{i=1}^{n} (v_i')^{-1}(v_i'(w_i)) = \sum_{i=1}^{n} w_i = 0\,.$$

We denote by (a_i, b_i) the open interval W_i and by (a, b) that of W. We have

$$a = a_i k_i \qquad \text{for at least one} \qquad i$$

and therefore

$$\lim_{x\to a} f_k(x) = \lim_{x\to a} \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{x}{k_i}\right) \geq \lim_{\substack{x\to a_i k_i \\ x>a}} (v_i')^{-1}\left(\frac{x}{k_i}\right) = \lim_{\substack{y\to a_i \\ y>a_i}} (v_i')^{-1}(y) = \infty\,.$$

Analoguously we get

$$\lim_{x\to b} f_k(x) = -\infty\,.$$

It follows that the continuous function f_k has range $\mathbb{R}$. Obviously f_k is differentiable on W with derivative

$$f_k'(x) = \sum_{i=1}^{n} \frac{1}{k_i} \frac{1}{v_i''\left((v_i')^{-1}\left(\frac{x}{k_i}\right)\right)} < 0\,. \qquad \text{QED}$$

LEMMA 2

Let w, $w := (w_1, \ldots, w_n) \in \mathbb{R}^n$, be a vector with $\sum_{i=1}^{n} w_i = 0$

and k, $k := (k_1, \ldots, k_n) \in \mathbb{R}^n$ the vector with $k_i := \frac{-1}{v_i'(w_i)} > 0\,.$

Furthermore let (*) and (**) respectively denote the system of equations (in $Y_1(z), \ldots, Y_n(z)$)

(*) $$z = f_k(k_i v_i'(Y_i(z))) \qquad i = 1, \ldots, n.$$

(**) $$\begin{cases} \sum_{i=1}^{n} Y_i(z) = z & \text{for all} \quad z \in \mathbb{R} \\ k_i v_i'(Y_i(z)) = k_1 v_1'(Y_1(z)) & \text{for} \quad i = 2, \ldots, n. \end{cases}$$

Then $Y(z) := (Y_1(z), \ldots, Y_n(z))$ with $Y_i(z) = (v_i')^{-1}\left(\frac{1}{k_i}(f_k)^{-1}(z)\right), i = 1, \ldots, n,$

is the unique solution of (*) resp. (**). The functions $Y_i(z)$, $i = 1, \ldots, n$, are strictly increasing and differentiable They satisfy $Y_i(0) = w_i$ for $i = 1, \ldots, n$

PROOF

From Lemma 1 it follows that $(f_k)^{-1}$ exists and is defined on $\mathbb{R}$. Therefore $Y(z)$ is well defined, strictly increasing and differentiable. By inverting equation (*) we see that $Y(z)$ is a solution and even the unique solution of (*) Obviously $Y(z)$ is also a solution of (**).

Note that

$$\sum_{i=1}^{n} Y_i(z) = f_k((f_k)^{-1}(z)) = z$$

$$k_i v_i'(Y_i(z)) = (f_k)^{-1}(z) = k_1 v_1'(Y_1(z)) \qquad \text{for} \qquad i = 2, \ldots, n.$$

Furthermore

$$Y_i(0) = (v_i')^{-1}\left(\frac{1}{k_i}(-1)\right) = w_i \qquad \text{(see proof of Lemma 1)}$$

Let $\tilde{Y}(z) := (\tilde{Y}_1(z), \ldots, \tilde{Y}_n(z))$ be another solution of (**) Then we have

$$z = \sum_{i=1}^{n} \tilde{Y}_i(z) = \sum_{i=1}^{n} (v_i')^{-1}\left(\frac{k_1}{k_i} v_1'(\tilde{Y}_1(z))\right)$$

$$= f_k(k_1 v_1'(\tilde{Y}_1(z))) = f_k(k_i v_i'(\tilde{Y}_i(z)))$$

But the solution of (*) is unique, so we have $\tilde{Y}(z) = Y(z)$ for all $z \in \mathbb{R}$. This completes the proof

QED

REFERENCES

BORCH, K (1960) The Safety Loading of Reinsurance Premiums *Skandinavisk Aktuarietidskrift* **43**, 163–184

BÜHLMANN, H (1980) An Economic Premium Principle *ASTIN Bulletin* **11**, 52–60

BÜHLMANN, H (1984) The General Economic Premium Principle *ASTIN Bulletin* **15**, 13 21

LIENHARD, M (1986) Calculation of Price Equilibria for Utility Functions of the HARA class *ASTIN Bulletin* **17**, 91–97

PRATT, J W (1964) Risk Aversion in the Small and in the Large *Econometrica* **32**, 122–136

ERICH WYLER
Anna-Heer-Strasse 28, CH-8057 Zürich, Switzerland.

FUZZY INSURANCE

BY JEAN LEMAIRE

Wharton School
University of Pennsylvania, U.S.A.

ABSTRACT

Fuzzy set theory is a recently developed field of mathematics, that introduces sets of objects whose boundaries are not sharply defined. Whereas in ordinary Boolean algebra an element is either contained or not contained in a given set, in fuzzy set theory the transition between membership and non-membership is gradual The theory aims at modelizing situations described in vague or imprecise terms, or situations that are too complex or ill-defined to be analysed by conventional methods This paper aims at presenting the basic concepts of the theory in an insurance framework. First the basic definitions of fuzzy logic are presented, and applied to provide a flexible definition of a "preferred policyholder" in life insurance. Next, fuzzy decision-making procedures are illustrated by a reinsurance application, and the theory of fuzzy numbers is extended to define fuzzy insurance premiums.

KEYWORDS

Fuzzy set theory; preferred policyholders in life insurance, optimal XL-retentions; net single premiums for pure endowment insurance

1 INTRODUCTION

In 1965, ZADEH published a paper entitled "Fuzzy Sets" in a little known journal, Information and Control, introducing for the first time sets of objects whose boundaries are not sharply defined. This paper gave rise to an enormous interest among researchers, and initiated the fulgurant growth of a new discipline of mathematics, fuzzy set theory. The number of papers related to the field exploded from 240 in 1975 (ZADEH et al.), to 760 in 1977 (GUPTA et al.), 2500 in 1980 (CHEN et al), and 5000 in 1987 (ZIMMERMAN). Today, there are many more researchers in fuzzy set theory than in actuarial science, and they form a much more international group, with important contributions from China, Japan, and the Soviet Union. Two monthy scientific journals publish new theoretical developments and applications, that are to be found in linguistics, risk analysis, artificial intelligence (approximate reasoning, expert systems), pattern analysis and classification (pattern recognition, clustering, image processing, computer vision), information processing, and decision-making. In this paper we will explore some possible applications of fuzzy set theory to insurance.

In ordinary Boolean algebra, an element is either contained or not contained in a given set. the transition from membership to non-membership is abrupt. Fuzzy sets, on the other hand, describe sets of elements or variables whose limits are ill-defined or imprecise The transition between membership and non-membership is gradual: an element can "more or less" belong to a set Consider for instance the set of "young drivers". In Boolean algebra, it is assumed that any individual either belongs or does not belong to the set of young drivers. This implies that the individual will move from the category of "young drivers" to the complementary set of "not young drivers" overnight Fuzzy set theory allows for grades of membership. Depending on the specific application, one might for instance decide that drivers under 20 are definitely young, that drivers over 30 are definitely not young, and that a 23-year-old driver is "more or less" young, or is young with a grade membership of 0.7, on a scale from 0 to 1

Fuzzy set theory thus aims at modelizing imprecise, vague, fuzzy information, which abound in real world situations. Indeed, many practical problems are extremely complex and ill-designed, hence difficult to modelize with precision To quote ZADEH, "as the complexity of a system increases, our ability to make precise and yet significant statements about its behaviour diminishes until a threshold is reached beyond which precision and significance become almost exclusive characteristics" Computers cannot adequately handle such problems, because machine intelligence still employs sequential (Boolean) logic. The superiority of the human brain results from its capacity of handling fuzzy statements and decisions, by adding to logic parallel and simultaneous information sources and thinking processes, and by filtering and selecting only those that are useful and relevant to its purposes. The human brain has many more thinking processes available and has developed a far greater filtering capacity than the machine A group of individuals is able to resolve the command "tall people in the back, short people in the front", a machine is not Fuzzy set theory explicitly introduces vagueness in the reasoning, hoping to provide decision-making procedures that are closer to the way the human brain performs.

A clear distinction has to be made between fuzzy sets and probability theory. Uncertainty should not be confused with imprecision Probabilities are primarily intended to represent a degree of knowledge about real entities, while the degrees of membership defining the strength of participation of an entity in a class are the representation of the degree by which a proposition is partially true Probability concepts are derived from considerations about the uncertainty of propositions about the real world Fuzzy concepts are closely related to the multivalued logic treatments of issues of imprecision in the definition of entities Hence, fuzzy set theory provides a better framework than probability theory for modelling problems that have some inherent imprecision The traditional approach to risk analysis, for instance, is based on the premise that probability theory provides the necessary and sufficient tools for dealing with the uncertainty and imprecision which underline the concept of risk in decision analysis The theory of fuzzy sets calls into question the validity of this premise. It does not equate imprecision with randomness It suggests that much of the uncertainty which is intrinsic in risk analysis is rooted in the fuzziness of the information which is resident in the data base and in the imprecision of the underlying probabilities. Classical probability theory has its effectiveness limited when dealing with problems in which some of the principal sources of uncertainty are non-statistical in nature

In the sequel we will present the basic principles of fuzzy logic, fuzzy decision-making, and fuzzy arithmetics, while developing three insurance examples We will show that fuzzy set theory could provide decision procedures that are much more flexible than those originating from conventional set theory Indeed, insurance executives and actuaries, much better trained to deal with uncertainty than with vagueness, have often transformed imprecise statements into "all-or-nothing" rules. For instance, Belgian insurers have used the fuzzy statistical evidence "Young drivers provoke more automobile accidents" to set up the a posteriori rating rule "Drivers under 23 years of age will pay a \$ 150 deductible if they provoke an accident". Hence "young" was equated with "under 23", a definite distorsion of the initial statement As another example, Belgian regulatory authorities define, for statistical purposes, a "severely wounded person" as "any person, wounded in an automobile accident, whose condition requires a hospital stay longer than 24 hours", a very arguable "de-fuzzification" of a fuzzy health condition

In Section 2 we will present the basic definitions of fuzzy logic and apply them to provide a more flexible definition of a "preferred policyholder" than the one currently used by some American life insurers Section 3 introduces the main concepts of fuzzy decision-making, and uses them to select an optimal Excess of Loss retention. Fuzzy arithmetics are presented in Section 4, and applied to compute the fuzzy premium of a pure endowment policy

First, let us introduce our three examples.

Problem 1 Definition of a preferred policyholder in life insurance

Heavy competition between American life insurers has resulted in a greater subdivison of policyholders than in Europe U.S. insurers first began, in the mid 1960s, to award substantial discounts to nonsmokers purchasing a term or a whole life insurance. Then the "preferred policyholder" category was further refined, and more discounts were granted to applicants who met very stringent health requirements, such as a cholesterol level not exceeding 200, a blood pressure not exceeding 130/80, . For instance, one company offers a non-smoker bonus of 65% more insurance coverage with no increase in premium if the applicant has not smoked for 12 months prior to application A bonus of 100% is offered if the applicant:

— has not smoked for the past 12 months, and
— has a resting pulse of 72 or below, and
— has a blood pressure that does not exceed 134/80, and
— has a total cholesterol reading not exceeding 200, and

— does not engage in hazardous sports, and
— rigorously follows a 3-times-a-week exercise program of at least 20 minutes, and
— is within specified height and weight limits, and
— has no more than one death in immediate family prior to 60 years of age due to kidney or heart disease, stroke or diabetes

Again this is a distorsion, or a least a very strict interpretation, of the medical statement "People who exercise, who do not smoke, who have a low level of cholesterol, low blood pressure, who are neither overweight nor severely underweight, ... have a higher life expectancy". Insurers demand all conditions to be strictly met, the slightest infringement leads to automatic rejection of the preferred category For instance, a cholesterol level of 201 implies that the preferred rates won't apply, even if the applicant meets all other requirements. A cholesterol level of 200 is accepted, a level of 201 is not! We will show that fuzzy set theory can be used to provide a more flexible definition of a preferred policyholder, that allows for some form of compensation between the selected criteria.

Problem 2. Selection of an optimal excess of loss retention

Imprecise statements seem to be pervasive in reinsurance practive, where vague recommendations and rules abound. "As a rule of thumb, an excess of loss (XL) retention should approximatively equal 1% of the premium income", "Our long-term relationship with our present reinsurer should in principle be maintained", "We could accept those conditions providing substantial retrocessions are offered", "A ball-park figure for the cost of this reinsurance program is $10 million", are fuzzy sentences frequently heard in practice. To illustrate fuzzy decision-making procedures, we shall consider the problem of the selection of the optimal retention of a pure XL treaty, given the four following fuzzy goals and constraints.

Goal 1: The ruin probability should be substantially decreased, ideally down to be neighbourhood of 10^{-5}.

Goal 2: The coefficient of variation of the retained portfolio should be reduced; if possible it should not exceed 3

Constraint 1. The reinsurance premium should not exceed 2 5% of the line's premium income by much.

Constraint 2. As a rule of thumb, the retention should approximatively be equal to 1% of the line's premium income

Problem 3 Computation of the fuzzy premium of a pure endowment policy

Forecasting interest rates is undoubtedly one of the most complex modelling problems. Money market interest rates seem to fluctuate according to monthly U.S. unemployment and trade deficit figures, vague statements made by Mr Kohl or Mr Greenspan, the markets' perception of Mr Bush's willingness to tackle the deficit problem, the mood of the participants to an OPEC meeting, etc. To compute insurance premiums over a 40-year span with a fixed interest rate of 4.75% then seems to be an exercise in futility. We will show that the introduction of fuzzy interest rates (and fuzzy survival probabilities) at least allows us to obtain a partial measure of our ignorance.

As illustrated by our examples, fuzzy set theory attempts to modelize imprecise expressions like "more or less young", "neither overweight nor underweight", "in the neighbourhood of", "in principle". In retreating from precision in the face of overpowering complexity, the theory explores the use of what might be called linguistic variables, that is, variables whose values are not numbers but words or sentences. In summary, fuzzy set theory endorses Bertrand Russell's opinion that

> "All traditional logic habitually assumes that precise symbols are being employed. It is therefore not applicable to this terrestrial life but only to an imagined celestial existence"

and rejects Yves Le Dantec's aphorism

> "That only is science which deals with the measurable".

2 FUZZY LOGIC AND FUZZY PREFERRED POLICYHOLDERS

2.1. Basic definitions

A fuzzy set is a class of objects in which there is no sharp boundary between those objects that belong to the class and those that do not. More precisely, let $X = \{x\}$ denote a collection of objects denoted generically by x A fuzzy set A in X is a set of ordered pairs

$$A = \{x, U_A(x)\}, \qquad x \in X$$

where $U_A(x)$ is termed the grade of membership of x in A, and $U_A : X \to M$ is a function from X to a space M, called the membership space Hence a fuzzy set A on a referential set X can be viewed as a mapping U_A from X to M. (Examples of membership functions are presented in all figures).

For our purposes it is sufficient to assume that M is the interval [0, 1], with 0 and 1 representing, respectively, the lowest and highest grade of membership The degree of membership of x in A corresponds to a "truth value" of the statement "x is a member of A". When M only contains the two points 0 and 1, A is nonfuzzy.

Problem 1

Let X be a set of prospective policyholders, $x = x(t_1, t_2, t_3, t_4)$. For simplicity, assume that the requirements for the status of "preferred policyholder" will be based on the values taken by 4 variables

t_1, the total level of cholesterol in the blood, in mg/dl,
t_2, the systolic blood pressure, in mm of Hg

t_3, the ratio (in %) of the effective weight to the recommended weight, as a function of height and build

t_4, the average consumption of cigarettes per day

Using a classical approach, an insurance company would for instance define a preferred policyholder as a nonsmoker with a cholesterol level that does not exceed 200, and a blood pressure that does not exceed 130, and a weight that is comprised between 85% and 110% of his recommended weight.

If a fuzzy set approach is to be used, membership functions have to be defined for all criteria.

National Institutes of Health nowadays recommend a level of less than 200 mg of cholesterol per deciliter of blood Levels between 200 and 240 mg/dl are considered to be borderline high The fuzzy set A of the people with a low level of cholesterol can then be defined by the membership function $U_A(x, t_1)$

$$U_A(x; t_1) = \begin{cases} 1 & t_1 \leqq 200 \\ 1-2\left(\dfrac{t_1-200}{40}\right)^2 & 200 < t_1 \leqq 220 \\ 2\left(\dfrac{240-t_1}{40}\right)^2 & 220 < t_1 \leqq 240 \\ 0 & 240 < t_1 \end{cases}$$

The normal systolic blood pressure is about 130 mm of mercury. People with a blood pressure greater than 170 are five times more likely to suffer from coronary heart disease than individuals with normal blood pressures Hence the fuzzy set B of the people with an acceptable blood pressure can be defined by the membership function $U_B(x, t_2)$

$$U_B(x, t_1) = \begin{cases} 1 & t_2 \leqq 130 \\ 1-2\left(\dfrac{t_2-130}{40}\right)^2 & 130 < t_2 \leqq 150 \\ 2\left(\dfrac{170-t_2}{40}\right)^2 & 150 < t_2 \leqq 170 \\ 0 & 170 < t_2 \end{cases}$$

Overweight and underweight people have a shorter life expectancy, skinniness being less primordial than obesity. This is reflected in the asymmetric membership function $U_C(x, t_3)$ that characterizes the fuzzy set C of the people with adequate weight

$$U_C(x, t_3) = \begin{cases} 0 & t_3 \leqq 60 \\ 2\left(\frac{t_3 - 60}{25}\right)^2 & 60 < t_3 \leqq 72.5 \\ 1 - 2\left(\frac{85 - t_3}{25}\right)^2 & 72.5 < t_3 \leqq 85 \\ 1 & 85 < t_3 \leqq 110 \\ 1 - 2\left(\frac{t_3 - 110}{20}\right)^2 & 110 < t_3 \leqq 120 \\ 2\left(\frac{130 - t_3}{20}\right)^2 & 120 < t_3 \leqq 130 \\ 0 & 130 < t_3 \end{cases}$$

Even light smokers are more prone to suffer from cancer and cardiovascular diseases than nonsmokers Hence they cannot be considered as "preferred" and the set D of the nonsmokers is nonfuzzy

$$U_D(x, t_4) = \begin{cases} 1 & t_4 = 0 \\ 0 & t_4 > 0. \end{cases}$$

The four selected membership functions are represented in Figure 1. Admittedly, there is some arbitrariness in the definiton of these membership functions, but fuzzy set theory contends that this is better than membership functions that abruptly jump from 1 to 0, in the classical approach

A fuzzy set is said to be normal iff $\text{Sup}_x\, U_A(x) = 1$ Subnormal fuzzy sets can be normalized by dividing each $U_A(x)$ by the factor $\text{Sup}_x\, U_A(x)$

$\bar{A}$ is said to be the complement of A iff $U_{\bar{A}}(x) = 1 - U_A(x) \quad \forall x$.

A fuzzy set is contained in or is a subset of a fuzzy set B ($A \subset B$) iff $U_A(x) \leqq U_B(x) \quad \forall x$.

The union of A and B, denoted $A \cup B$, is defined as the smallest fuzzy set containing both A and B Its membership function is given by

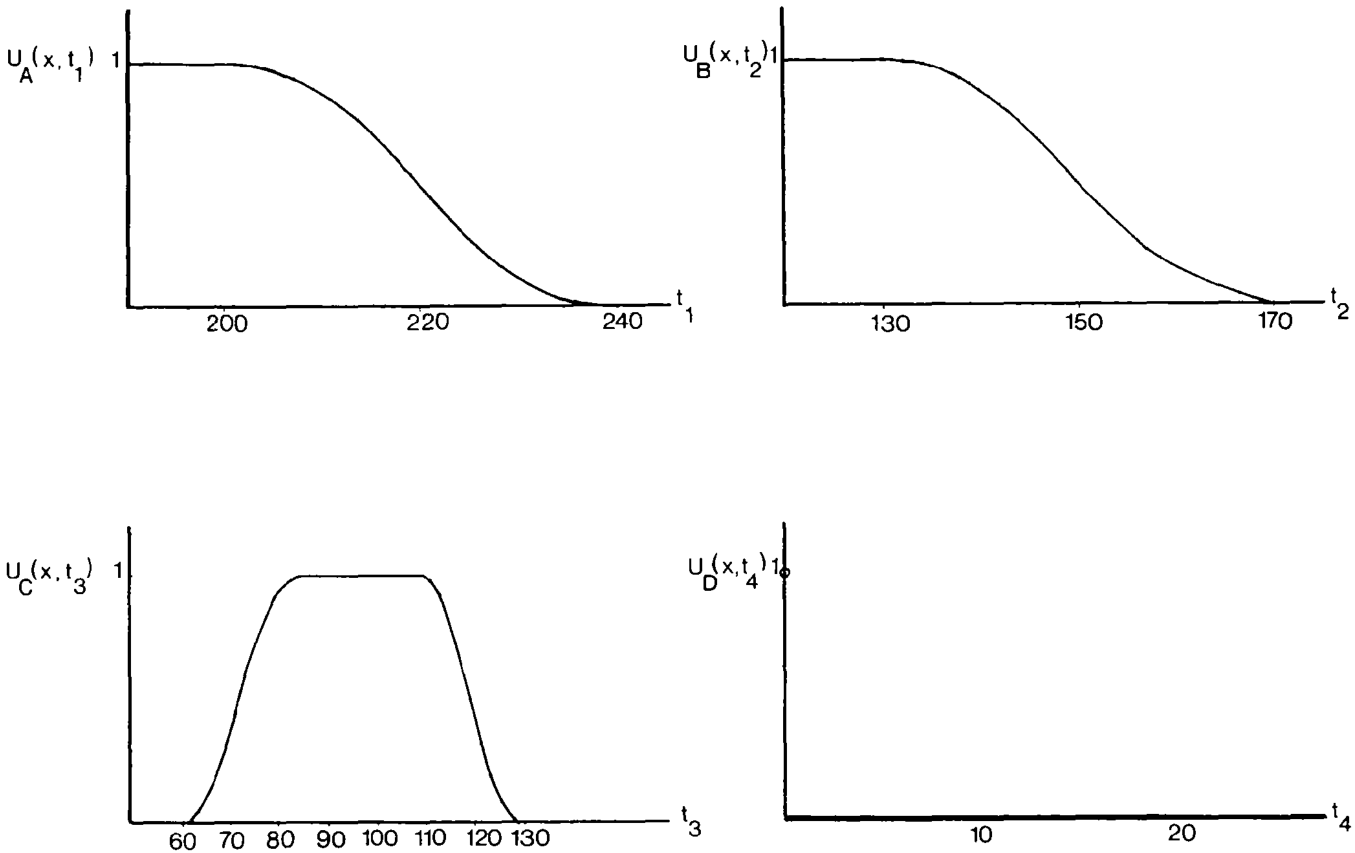
$$U_{A \cup B}(x) = \max\,[U_A(x), U_B(x)] \qquad x \in X$$

The intersection of A and B, denoted $A \cap B$, is defined as the largest fuzzy set contained in both A and B. Its membership function is given by

$$U_{A \cap B}(x) = \min\,[U_A(x), U_B(x)] \qquad x \in X$$

The notion of intersection bears a close relation to the notion of the connective "and", just as the union of A and B bears a close relation to the connective "or". It can be shown that these definitions of fuzzy union and intersection are

FIGURE 1 Membership functions Problem 1



the only ones that naturally extend the corresponding standard set theory notions, by satisfying all the usual requirements of associativity, commutativity, idempotency and distributivity.

Problem 1

The fuzzy set E of the nonsmoking individuals with low cholesterol, acceptable blood pressure and adequate weight is the intersection of the 3 fuzzy sets A, B, C, and the nonfuzzy set D. Its membership function is given by

$$U_E(x; t_1, t_2, t_3, t_4) = \min [U_A(x; t_1), U_B(x, t_2), U_C(x; t_3), U_D(x; t_4)]$$

So an individual can only be a full member of E if he doesn't smoke, has a cholesterol level not exceeding 200, a blood pressure not above 130, and a weight no less than 85% and no more than 110% of his recommended or ideal weight. This corresponds to the classical approach.

A nonsmoker $x = x(210, 145, 112, 0)$ with a cholesterol level of 210, a blood pressure of 145, and who is overweight by 12% is a member of E with a grade of membership

$$U_E(x, 210, 145, 112, 0) = \min (0.875, 0.71875, 0\ 98, 1) = 0.71875.$$

In other words, the "$\cap$" operation assigns a grade of membership that corresponds to the most severe of the infringements to "perfection", in this case blood pressure. Cumulative effects and interactions between the criteria are ignored, which is not realistic. Obviously, the health consequences of high blood pressure are worse when there is also an excess of weight and cholesterol. Also, since only the most severe condition is considered, it is impossible to introduce compensations or trade-offs in decision rules. A mild excess of weight cannot be compensated by ideal cholesterol and blood pressure

2.2. Other definitions of the intersection

The minimum operator that characterizes the intersection corresponds to the "logical and" Other definitions of the intersection have been suggested, they correspond to "softer", more flexible interpretations of the connective "and". They all amount to exactly the same in the conventional case of degrees of membership restricted to 0 and 1. The selection of a specific operator will depend on its possibilities to allow for cumulative effects, interactions, and compensations between the criteria. We wish the following properties to be satisfied.

Property 1 (cumulative effects): Two infringements are worse than one.

$$U_{A\cap B}(x) < \min [U_A(x), U_B(x)] \quad \text{if} \quad U_A(x) < 1 \quad \text{and} \quad U_B(x) < 1 .$$

Property 2 (interactions between criteria). Assume $U_A(x) < U_B(x) < 1$. Then the effect of a decrease of $U_A(x)$ on $U_{A\cap B}(x)$ may depend on $U_B(x)$

Property 3 (compensations between criteria): If $U_A(x)$ and $U_B(x) < 1$, the effect of a decrease of $U_A(x)$ on $U_{A \cap b}(x)$ can be erased by an increase of $U_B(x)$ (unless, of course, $U_B(x)$ reaches 1).

The algebraic product F of A and B is denoted AB and is defined by

$$U_{AB}(x) = U_A(x) \cdot U_B(x)$$

The bounded difference G of A and B is denoted $A \ominus B$ and is defined by

$$U_{A \ominus B}(x) = \max [0, U_A(x) + U_B(x) - 1]$$

The Hamacher operator H defines the intersection of two fuzzy sets A and B by

$$U_H^p(x) = \frac{U_A(x) \cdot U_B(x)}{p + (1-p)\,[U_A(x) + U_B(x) - U_A(x)\, U_B(x)]} \qquad 0 \leqq p \leqq 1$$

The Yager operator Y defines the intersection of two fuzzy sets A and B by

$$U_Y^p(x) = 1 - \min \{1, [(1 - U_A(x))^p + (1 - U_B(x))^p]^{1/p}\} \qquad p \geqq 1$$

Problem 1

The generalized operators provide a more realistic way of modelling this specific problem because they explicitly allow for compensations and interactions between the selected criteria First consider the algebraic product. The grade of membership of individual $x(210, 145, 112, 0)$ in the fuzzy set $F = ABCD$ is

$$U_F(x; 210, 145, 112, 0) = (0\ 875)\,(0.71875)\,(0.98)\,(1) = 0\ 6163$$

The effect of high blood pressure is here amplified by the presence of a slight obesity and a cholesterol level mildly above normal This operator satisfies all three properties.

The grade of membership of the same individual in the fuzzy set $G = A \ominus B \ominus C \ominus D$ corresponding to the bounded difference operation is

$$U_G(x; 210, 145, 112, 0) = \max [0, 0\ 875 + 0.71875 + 0\ 98 + 1 - 3] = 0\ 57375$$

Hence the effects of the criteria are additive; no interactions are introduced, since the consequences of cholesterol are the same whatever the blood pressure and the weight. This operator satisfies properties 1 and 3, but not property 2

The minimum and algebraic product operators model two extreme situations. The minimum operator does not satisfy any property Compensations and interactions cannot be introduced. The algebraic product allows for compensation and maximum interaction, since the effect of one criterion fully impacts the others. The Hamacher and Yager operators model intermediate situations, with flexibility provided by the parameter p.

The Hamacher operator reduces to the algebraic product when $p = 1$. For $p < 1$, the denominator is less than 1 and $U_H(x) > U_F(x)$: the product operator is "softened"; this operator models weaker interactions It reduces the effect of combined infringements The reduction effect is greater for severe infringements. Also, the lower the selected p, the greater the reduction effect Hence this operator can be used if it is considered that the combined effect of high cholesterol and high blood pressure is somewhat less than multiplicative Selecting $p = 0.5$ for our example, we obtain successively

$$U_H^{1/2}(x, 210, 145) = \frac{(0\,875)(0\,71875)}{0\,5+(1-0\,5)[0.875+0.71875-(0.875)(0.71875)]} = 0.6402$$

$$U_H^{1/2}(x, 210, 145, 112, 0) = U_H^{1/2}(x, 210, 145, 112)$$

$$= \frac{(0\,6402)(0\,98)}{0.5+(1-0\,5)[0\,6402+0\,98-(0\,6402)(0.98)]} = 0\,6296$$

This operator satisfies all three properties.

The Yager operator reduces to the bounded difference operator when $p = 1$, and to the minimum operator when $p \to \infty$. $U_Y^p(x)$ is an increasing function of p. Hence all intermediate situations can be modelled, from the strongest to the weakest "and" Selecting $p = 2$, we obtain

$$U_Y^2(x) = 1-\min\{1, [(1-0.875)^2+(1-0\,71875)^2+(1-0\,98)^2]^{1/2}\} = 0.69157$$

This operator satisfies all three properties, except in the case $p = \infty$.

2.3. Selection of a decision rule

If A is a fuzzy subset of X, its a-cut A_a is defined as the nonfuzzy subset such that

$$A_a = \{x | U_a(x) \geqq a\} \quad \text{for} \quad 0 < a \leqq 1$$

An a-cut can be interpreted as an error interval whose truth value is a.

Problem 1

The notion of a-cut provides a flexible way of defining preferred policyholders. The "classical" approach corresponds to 1-cuts such as E_1 or F_1. Lower values of a provide generalizations of this definition For instance preferrred customers could be defined as the members of $E_{0\,75}$ or $F_{0\,60}$. $E_{0\,75}$ is the set of policyholders for which the grade of membership attains at least 0 75 for each of the selected criteria (for our specific membership functions, $t_1 \leqq 214$, $t_2 \leqq 144$, $76.2 < t_3 \leqq 117.1$, $t_4 = 0$). Hence this amounts to relaxing all criteria in a uniform way

$F_{0\,60}$ is the set of policyholders for which the product of the four grades of membership attains at least 0 60 The latter definition is more realistic because it allows for interactions and compensations An excess of blood pressure can for instance be compensated by normal or near-normal weight and cholesterol levels Policyholder $x(210, 145, 112, 0)$ is accepted as preferred using the second criterion. He is not accepted if the first criterion is used

Similar decision rules can be constructed using the other operators, if medical considerations hint that they provide a better model of the problem.

2.4. Fuzzy operations

The concept of grades of membership allows to define the following operations that have no counterpart in ordinary set theory; they are uniquely fuzzy.

Concentration: A fuzzy set is concentrated by reducing the grade of membership of all elements that are only partly in the set, in such a way that the less an element is in the set, the more its grade of membership is reduced The concentration of a fuzzy set A is denoted CON (A) and defined by

$$U_{\mathrm{CON}(A)}(x) = U_A^a(x) \qquad a > 1$$

Dilation: Dilation is the opposite of concentration A fuzzy set is dilated or stretched by increasing the grade of membership of all elements that are partly in the set. The dilation of a fuzzy set A is denoted DIL (A) and defined by

$$U_{\mathrm{DIL}(A)}(x) = U_A^a(x) \qquad a < 1$$

a is called the power of the operation.

Intensification: A fuzzy set can be intensified by increasing the grade of membership of all the elements that are at least half in the set and decreasing the grade of membership of the elements that are less than half in the set The intensification of a fuzzy set is denoted INT (A) and is defined by

$$U_{\mathrm{INT}(A)}(x) = \begin{cases} 2\,U_A^2(x) & 0 < U(x) \leqq 0\ 5 \\ 1-2[1-U_A(x)]^2 & 0\ 5 < U(x) \leqq 1 \end{cases}$$

Fuzzification. A fuzzy set can be fuzzified or de-intensifed by increasing the extent of its fuzziness. There are several ways of achieving this.

Problem 1

The operations of concentration and dilation roughly approximate the effect of the linguistic modifiers "very" and "more or less". They are used whenever the different criteria have to be weighted. The presentation of problem 1 so far implicitly assumes that each criterion has the same importance. If for medical reasons this is not desirable, fuzzy operations can be used. Suppose that cholesterol level is the better predictor of future heart problems, while the importance of blood pressure has to be downgraded. This can be reflected by assigning powers of 2 and 0.5 to the two criteria. The modified fuzzy set $\tilde{E}$, corresponding to the minimum operator, is characterized by

$$U_{\tilde{E}}(x; t_1, t_2, t_3, t_4) = \min [U_A^2(x, t_1), U_B^{1/2}(x; t_2), U_C(x; t_3), U_D(x; t_4)]$$

The modified fuzzy set $\tilde{F}$, corresponding to the algebraic product, has the membership function

$$U_{\tilde{F}}(x; t_1, t_2, t_3, t_4) = U_A^2(x, t_1)\, U_B^{1/2}(x; t_2)\, U_C(x; t_3)\, U_D(x; t_4)$$

Prospective policyholder $x(210, 145, 112, 0)$ has a grade of membership of

$$\min [(0.875)^2, (0.71875)^{1/2}, 0.98, 1] = 0.7656$$

in $\tilde{E}$, and of

$$(0.875)^2 \cdot (0.71875)^{1/2}\ (0.98)\ (1) = 0.6361$$

in $\tilde{F}$. He is now accepted as a preferred customer under each of the two criteria of Section 2 3, since $x(210, 145, 112, 0)$ is included in both $\tilde{E}_{0.75}$ and $\tilde{F}_{0.60}$.

3 DECISION-MAKING WITH FUZZY GOALS AND CONSTRAINTS AND FUZZY REINSURANCE

In the classical approach to decision-making, the principal ingredients of a decision problem are (*a*) a set of alternatives, (*b*) a set of constraints on the choice between different alternatives, and (*c*) an objective function which associates with each alternative its evaluation. There is however an intrinsic similarity between objective functions and constraints, a similarity that becomes apparent when for instance Lagrangian multipliers are introduced

This similarity is made explicit in the formulation of a decision problem in a fuzzy environment Let $X = \{x\}$ be a given set of alternatives. A fuzzy goal G in X, or simply a goal G, is expressed and identified with a given fuzzy set G in X In other words, a fuzzy goal is an objective which can be characterized as a fuzzy set in the space of alternatives In the classical approach, the objective function serves to define a linear ordering on the set of alternatives. Clearly the membership function $U_G(x)$ of a fuzzy goal serves the same purpose, and may even be derived from a given objective function by normalization, which leaves the linear ordering unaltered Such normalization provides a common denominator for the various goals and constraints and makes it possible to treat them alike. A fuzzy constraint C in X, or simply a constraint C, is similarly defined to be a fuzzy set C in X. An important aspect of those definitions is thus that the notions of goal and constraint both are defined as fuzzy sets in the space of alternatives. Hence they can be treated identically in the decision process Since we want to satisfy (optimize) the objective function as well as the constraints, a decision in a fuzzy environment is defined as the selection of activities which simultaneously satisfy objective functions and constraints. A decision can therefore be viewed as the intersection of fuzzy constraints and fuzzy objective function(s) The relationship between constraints and objective functions in a fuzzy environment is therefore fully symmetric

Assume we are given a finite set of alternatives $X = \{x_1, x_2, \ldots, x_n\}$, a set of goals $G_1, \ldots, G_p$, characterized by their respective membership functions $U_{G_1}(x), \ldots, U_{G_p}(x)$, and a set of constraints $C_1, \ldots, C_q$, characterized by their respective membership functions $U_{C_1}(x), \ldots, U_{C_q}(x)$. Finiteness is assumed for expository purposes only and can be easily relaxed.

A decision is a choice or a set of choices drawn from the available alternatives, satisfying the constraints and the goals. The constraints and goals combine to form a decision D, which is naturally defined as the intersection of the fuzzy sets G's and C's.

$$D = G_1 \cap G_2 \cap \ldots \cap G_p \cap C_1 \cap C_2 \cap \ldots \cap C_q$$

Consequently a decision D is a fuzzy set in the space of alternatives whose membership function is

$$U_D(x) = \min\,[U_{G_1}(x), \ldots, U_{G_p}(x), U_{C_1}(x), \ldots, U_{C_q}(x)]$$

This decision membership function can be interpreted as the degree to which each of the alternatives satisfies the goals and constraints As in example 1, concentrations and dilations can be performed to reflect unequal importances of the goals and constraints, and other intersection operators can be used.

Let K be the (nonfuzzy) set consisting of all the alternatives for which $U_D(x)$ reaches its maximal value K is called the optimizing set, and any alternative in K is an optimal decision. The decision-maker simply selects as best alternative the one that has the maximum value of membership in D

This decision-making procedure is essentially a maximin technique, similar to the selection of an optimal strategy in noncooperative game theory. For each alternative the minimum possible grade of membership of all the goals and constraints is computed to obtain D Then the maximum value over the alternatives in D is selected

Problem 2

Given the formulation of the problem, a reinsurance program is characterized by its XL deductible, and evaluated by means of 4 different variables

t_1 = probability of ruin ($\times 10^4$)

t_2 = coefficient of variation of the retained portfolio

$$t_3 = \frac{\text{reinsurance premium}}{\text{cedent's premium income}} \quad (\text{in } \%)$$

$$t_4 = \frac{\text{deductible}}{\text{cedent's premium income}} \quad (\text{in } \%)$$

Assume the reinsurer offers 10 different XL deductibles, arranged in increasing order ($x = 1, 2, \ldots, 10$). The values taken by the selected variables are provided in Table 1

TABLE 1

CHARACTERISTICS OF THE 10 XL REINSURANCE PROGRAMS

Program		1	2	3	4	5	6	7	8	9	10
G_1	t_1	339	280	200	200	313	339	360	388	419	465
G_2	t_2	2.98	3.00	3.03	3.07	3.12	3.19	3.28	3.52	3.80	4.20
C_1	t_3	3.20	3.00	2.85	2.73	2.64	2.57	2.52	2.48	2.45	2.43
C_2	t_4	4	6	8	9	1.0	1.1	1.2	1.4	1.6	1.8

The following membership functions have been chosen. They are represented in Figure 2.

Goal 1 (probability of ruin)

$$U_{G_1}(x, t_1) = \begin{cases} 1 & t_1 \leqq .00002 \\ 1 - 2\left(\dfrac{t_1 - .00002}{.00008}\right)^2 & .00002 < t_1 \leqq .00006 \\ 2\left(\dfrac{.0001 - t_1}{.00008}\right)^2 & .00006 < t_1 \leqq .0001 \\ 0 & .0001 < t_1 \end{cases}$$

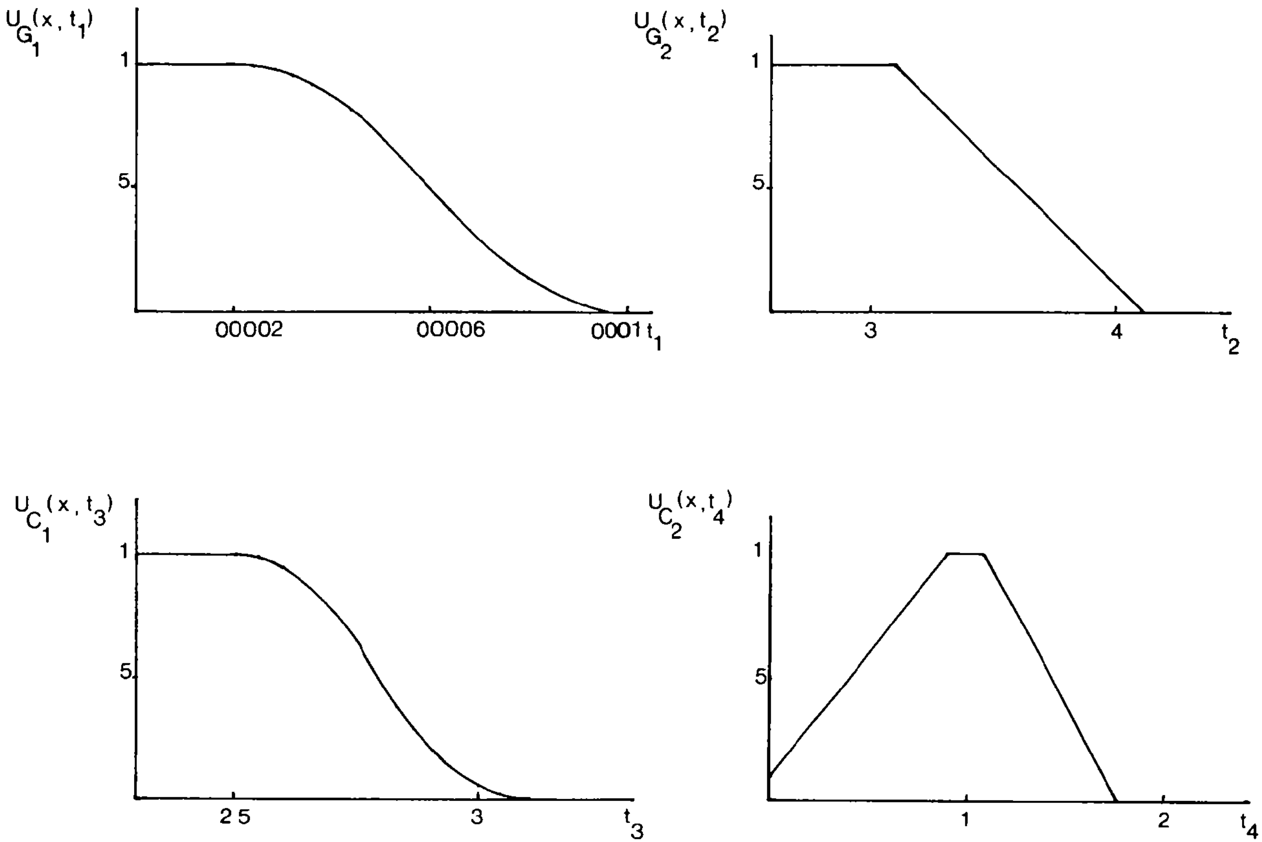
Goal 2 (coefficient of variation)

$$U_{G_2}(x, t_2) = \begin{cases} 1 & t_2 \leqq 3.1 \\ 4.1 - t_2 & 3.1 < t_2 \leqq 4.1 \\ 0 & 4.1 < t_2 \end{cases}$$

Constraint 1 (reinsurance premium)

$$U_{C_1}(x, t_3) = \begin{cases} 1 & t_3 \leqq 2.5 \\ 1 - 2\left(\dfrac{t_3 - 2.5}{0.6}\right)^2 & 2.5 < t_3 \leqq 2.8 \\ 2\left(\dfrac{3.1 - t_3}{0.6}\right)^2 & 2.8 < t_3 \leqq 3.1 \\ 0 & 3.1 < t_3 \end{cases}$$

FIGURE 2 Membership functions Problem 2



Constant 2 (deductible)

$$U_{C_2}(x; t_4) = \begin{cases} t_4 + 0.1 & 0 < t_4 \leqq 0.9 \\ 1 & 0.9 < t_4 \leqq 1.1 \\ 2.65 - 1.5\,t_4 & 1.1 < t_4 \leqq 1.7667 \\ 0 & 1.7667 < t_4 \end{cases}$$

Given those membership functions, the grades of membership for all alternatives are easily computed. They are presented in Table 2.

TABLE 2

GRADES OF MEMBERSHIP OF THE 10 DIFFERENT PROGRAMS

Program	1	2	3	4	5	6	7	8	9	10
G_1	94	98	1	1	96	94	92	89	85	78
G_2	1	1	1	1	1	91	82	58	30	0
C_1	0	0	06	35	71	89	97	998	1	1
C_2	5	7	9	1	1	1	85	55	25	0

The membership function $U_D(x)$ of the decision D is obtained by simply taking the minimum of the U's, for each alternative, as shown in Table 3

TABLE 3

MEMBERSHIP FUNCTION OF D

Program	1	2	3	4	5	6	7	8	9	10
$U_D(x)$	0	0	06	35	71	89	82	55	25	0

Note that no alternative has full membership in D: fuzzy set D is subnormal This of course reflects the fact that the specified goals and constraints conflict with one another, ruling out the existence of an alternative which fully satisfies all of them

In our case, when all goals and constraints are considered to be of equal importance, the ruin probability criterion is inoperative, it does not influence the decision. The membership function of D is based on the first constraint for alternatives 1 to 6, on the second goal for alternative 7, and on the second constraint for alternatives 8 to 10

The optimal decision is program 6, corresponding to a retention of 1.1 % of the cedent's premium income. This alternative fully satisfies the second constraint, given our selection of membership functions. The other constraint and the two goals are conflicting and cannot be fully satisfied The worst infringement is the reinsurance premium, considered to be too high.

Assume now that, after reviewing the preceding analysis, the manager of the reinsurance department decides that the first constraint C_1 is of paramount importance, and accordingly assigns it a higher weight. A concentration of the fuzzy set C_1, with $a = 2$, is then performed: the values of $U_{C_1}(x, t_3)$ are simply squared. This has the effect of decreasing the membership function of that important constraint and making it more influential in the determination of D. It is easily seen that the optimal decision becomes program 7. This illustrates an inherent weakness of fuzzy decision-making the sensitivity of the optimal solution to the particular selection of membership functions. And it is difficult to avoid an important element of subjectivity in the determination of those functions (see, however, CIVANLAR and TRUSSEL (1986) and DISHKANT (1981) for attempts to construct membership functions using statistical data)

The preceding analysis used the "hard" definition of the connective "and", since the minimum operator was used as intersection. As illustrated in Example 1, this excludes all forms of compensations and interactions between the goals and constraints In some managerial problemes the decision maker might wish to be less restrictive. For instance, he might not really want to actually maximize the objective function, but rather reach some aspiration level, which might not even be definable crisply (his objective might be to "improve the present cost situation considerably", for instance). Or the "$\leqq$" sign in a constraint might not be meant in the strict mathematical sense, but small violations might be acceptable, especially if an important improvement in the objective function results (effective expenditures might slightly exceed a budget constraint, for instance) Hence in many cases it is more appropriate to use a "softer" aggregation operator than the minimum, like the bounded difference or the Yager operator. A decision is then defined as the confluence of goals and constraints

$$U_D(x) = U_{G_1}(x) * \ldots * U_{G_p}(x) * U_{C_1}(x) \quad * U_{C_q}(x),$$

where $*$ is the selected operator.

It is easily checked, for instance, that if the algebraic product is used instead of the minimum operator, program 6 is the optimal solution of problem 2, with program 5 a close second.

4. FUZZY ARITHMETICS AND FUZZY INSURANCE PREMIUMS

DEFINITIONS. A fuzzy number is a fuzzy subset of the real line whose highest membership values are clustered around a given real number. The membership function is monotonic on both sides of this real number More precisely, a fuzzy number A is a fuzzy subset of the real line R whose membership function $U_A(x) = U_A(x; a_1, a_2, a_3, a_4)$ is:

(i) a continuous mapping from R to the closed interval [0, 1]
(ii) zero on the interval $(-\infty, a_1]$
(iii) strictly increasing on the interval $[a_1, a_2]$
(iv) one on the interval $[a_2, a_3]$
(v) strictly decreasing on the interval $[a_3, a_4]$
(vi) zero on the interval $[a_4, \infty)$,

where $a_1 < a_2 < a_3 < a_4$. (Examples of membership functions of fuzzy numbers are presented in Figure 3). The increasing part of $U_A(x)$, on interval $[a_1, a_2]$, is denoted $U_{A1}(x)$, the decreasing part of $U_A(x)$, on interval $[a_3, a_4]$, is denoted $U_{A2}(x)$. Alternatively, the inverse functions of $U_{A1}(x)$ and $U_{A2}(x)$, $U_{A1}^{-1}(y)$ and $U_{A2}^{-1}(y)$ can be used; they are denoted $V_{A1}(y)$ and $V_{A2}(y)$.

If $a_1 = a_2 = a_3 = a_4$, A is an ordinary real number.

A fuzzy number A is said to be positive if $a_1 > 0$. It is negative if $a_4 < 0$

Let A and B be two fuzzy numbers with membership functions $U_A(x) = U_A(x; a_1, a_2, a_3, a_4)$ and $U_B(x) = U_B(x; b_1, b_2, b_3, b_4)$. The membership function of the sum C of A and B, denoted $A \oplus B$, is defined as

$$U_C(z) = \max_{x+y=z} \min [U_A(x), U_B(y)] \qquad (x, y, z) \in R^3$$

$$= \max \min [U_A(x), U_B(z-x)].$$

It can be shown (see for instance DUBOIS and PRADE (1978) and (1980)) that the sum of fuzzy numbers is associative and commutative, and that

(i) $U_C(z) = 0 \qquad z \in (-\infty, a_1+b_1]\, U[a_4+b_4, \infty]$

(ii) $U_C(z)$ is strictly increasing in $[a_1+b_1, a_2+b_2]$, and strictly decreasing in $[a_3+b_3, a_4+b_4]$

(iii) $U_C(z) = 1 \qquad z \in [a_2+b_2, a_3+b_3]$

(iv) $U_{C1}(z) = [U_{A1}^{-1}(z) + U_{B1}^{-1}(z)]^{-1}$ or $V_{C1}(z) = V_{A1}(z) + V_{B1}(z)$

$U_{C2}(z) = [U_{A2}^{-1}(z) + U_{B2}^{-1}(z)]^{-1}$ or $V_{C2}(z) = V_{A2}(z) + V_{B2}(z)$.

The product D of A and B, denoted $A \odot B$, is defined by

$$U_D(z) = \max_{xy=z} \min [U_A(x), U_B(y)] \qquad (\text{assuming } a_1, b_1 > 0)$$

It can be shown that D is a fuzzy number, with $d_1 = a_1 b_1$, $d_2 = a_2 b_2$, $d_3 = a_3 b_3$, $d_4 = a_4 b_4$,

$U_{D1}(z) = [U_{A1}^{-1}(z)\; U_{B1}^{-1}(z)]^{-1}$ or $V_{D1}(z) = V_{A1}(z)\; V_{B1}(z)$

$U_{D2}(z) = [U_{A2}^{-1}(z)\; U_{B2}^{-1}(z)]^{-1}$ or $V_{D2}(z) = V_{A2}(z)\; V_{B2}(z)$.

The product is associative and commutative, and distributive on $\oplus$. The n^{th} power of A is naturally recursively defined as

$$A^n = A \odot A^{n-1}$$

The only reference dealing with finance applications of fuzzy arithmetics seems to be BUCKLEY (1987), who defined the fuzzy extensions of the notions of present and accumulated value, and annuities, and showed how to compare fuzzy cash flows by means of extended net present value and internal rate of return methods. Problem 3 is a straightforward generalization of that paper to an insurance problem

Problem 3

Let us compute the net single premium of a $ 1000, 10-year pure endowment policy, on a life aged (55), where $p = {}_{10}p_{55}$ is 0.87. The interest rate i is fuzzy and assumed to be approximately equal to 6%, as modelized by

$$U_i(x) = \begin{cases} 0 & x \leqq 1.03 \\ U_{i1}(x) = 50x - 51.5 & 1.03 < x \leqq 1.05 \\ 1 & 1.05 < x \leqq 1.07 \\ U_{i2}(x) = 54.5 - 50x & 1.07 < x \leqq 1.09 \\ 0 & 1.09 < x \end{cases}$$

(see Figure 3, upper left). As shown by the definitions of $\oplus$ and $\odot$, it is easier to use the inverse functions

$$V_{i1}(y) = 1.03 + 0.02y \quad \text{and} \quad V_{i2}(y) = 1.09 - 0.02y .$$

The present value $PV(S, n)$ of a positive fuzzy amount S, n periods in the future, if the fuzzy interest rate is i per period, can be defined as

$$PV(S, n) = S \odot (1 \oplus i)^{-n}$$

This definition makes sense given the associativity and the distributivity properties of $\odot$ Note however that, generally, $PV(S, n) \odot (1 \oplus i)^n$ will not be equal to S. Since the face value and the survival probability are nonfuzzy, the single fuzzy premium A of the policy,

$$A = 1000 \cdot 0.87 \cdot (1 \oplus i)^{-10} ,$$

is defined by the membership function

$$U_A(x) = \begin{cases} 0 & x \leqq 367.50 \\ U_{A1}(x) \text{ or } V_{A1}(y) & 367.50 < x \leqq 442.26 \\ 1 & 442.26 < x \leqq 534.10 \\ U_{A2}(x) \text{ or } V_{A2}(y) & 534.10 < x \leqq 647.36 \\ 0 & 647.36 < x \end{cases}$$

where

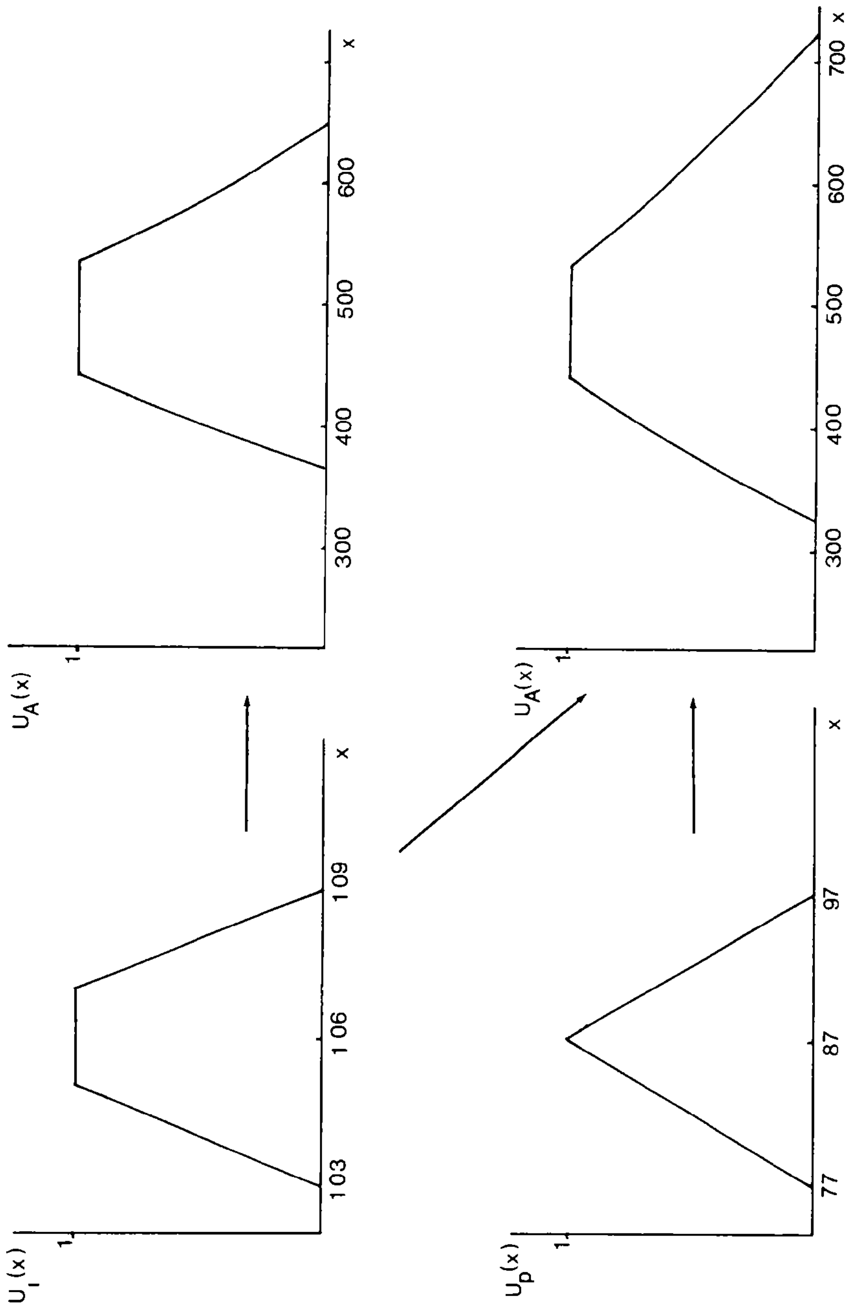
$$V_{A1}(y) = 870(1.09 - 0.02y)^{-10}$$

and

$$V_{A2}(y) = 870(1.03 + 0.02y)^{-10} \qquad (0 \leqq y \leqq 1)$$

This function is represented in Figure 3, upper right

FIGURE 3 Membership functions Problem 3



Next assume that $p = {}_{10}p_{55}$ is also fuzzy, with membership function

$$U_p(x) = \begin{cases} 0 & (x \leqq 0.77)\, U(x > 0.97) \\ 10x - 7.7 & 0.77 < x \leqq 0.87 \\ 9.7 - 10x & 0.87 < x \leqq 0.97 \end{cases}$$

and inverse functions $V_{p1}(y) = 0.77 + 0.01y$ and $V_{p2}(y) = 0.97 - 0.01y$ (see Figure 3, lower left).

The membership function of the premium A now becomes

$$U_A(x) = \begin{cases} 0 & x \leqq 325.26 \\ U_{A1}(x) \text{ or } V_{A1}(y) & 325.26 < x \leqq 442.26 \\ 1 & 442.26 < x \leqq 534.10 \\ U_{A2}(x) \text{ or } V_{A2}(y) & 534.10 < x \leqq 721.77 \\ 0 & 721.77 < x \end{cases}$$

where

$$V_{Aj}(y) = 1000 \cdot V_{pj}(y) \cdot [1 + V_{i,3-j}(y)]^{-10} \qquad j = 1, 2$$

$$V_{A1}(y) = 1000(0.77 + 0.1y)(1.09 - 0.02y)^{-10}$$

$$V_{A2}(y) = 1000(0.97 - 0.1y)(1.03 + 0.02y)^{-10}$$

This membership function, represented in the lower right part of Figure 3, reflects the increased fuzziness.

It is also possible (see BUCKLEY (1987)) to fuzzify the number of periods n.

5 FUZZY SETS LITERATURE

The literature about fuzzy sets is abundant and highly specialized. A good introductory textbook is ZIMMERMANN (1987), despite the important number of misprints. More specialized textbooks are KAUFMANN (1975) and DUBOIS and PRADE (1980). The seminal papers about fuzzy decision-making are BELLMAN and ZADEH (1970) and YAGER and BESSON (1976). Fuzzy graph theory, fuzzy linear and dynamic programming and extensions of other operations research methods are surveyed in ZIMMERMANN (1985) and (1987). Reference papers for applications of fuzzy set theory to statistics are HESHMATY and KANDER (1985), BUCKLEY (1985) and JAJUGA (1986). Topics of interest for actuaries where fuzzy applications have been developed include game theory (AUBIN (1981), BUTNARIU (1978, 1980)), economics (CHANG (1977), CHEN et al (1980)), and utility theory (MATHIEU-NICOT (1986))

REFERENCES

AUBIN, J P (1981) Cooperative Fuzzy Games *Math Oper Res* **6**, 1 13
BELLMAN, R and ZADEH, L (1970) Decision-making in a Fuzzy Environment *Management Science* **17**, B141 B164
BUCKLEY, J (1985) Fuzzy Decision Making with Data Applications to Statistics *Fuzzy Sets and Systems* **16**, 139 147
BUCKLEY, J (1987) The Fuzzy Mathematics of Finance *Fuzzy Sets and Systems* **21**, 257 273
BUTNARIU, D (1978) Fuzzy Games A Description of the Concepts *Fuzzy Sets and Systems* **1**, 181 192
BUTNARIU, D (1980) Stability and Shapley Value for an *n*-persons Fuzzy Game *Fuzzy Sets and Systems* **4**, 63 72
CHANG, S (1977) Application of Fuzzy Set Theory to Economics *Kybernetics* **6**, 203 207
CHEN, G, LEE, S and YU, E (1980) Application of Fuzzy Set Theory to Economics In *Advances in Fuzzy Sets, Possibility Theory, and Applications*, (ed P Wang) pp 227 305, Plenum Press, New York
CIVANLAR, R and TRUSSELL, J (1986) Constructing Membership Functions Using Statistical Data *Fuzzy Sets and Systems* **18**, 1 13
DISHKANT, H (1981) About Membership Functions Estimations *Fuzzy Sets and Systems* **5**, 141-147
DUBOIS, D and PRADE, H (1978) Operations on Fuzzy Numbers *Int J Systems Sci* **9**, 613-626
DUBOIS, D and PRADE, H (1980) *Fuzzy Sets and Systems* Academic Press New York
GUPTA, M, SARIDIS, G and GAINES, B, Editors (1977) *Fuzzy Automata and Decision Processes* North Holland, Amsterdam
HESHMATY, B and KANDER, A (1985) Fuzzy Linear Regression and its Applications to Forecasting in Uncertain Environment *Fuzzy Sets and Systems* **15**, 159 191
JAJUGA, K (1986) Linear Fuzzy Regression *Fuzzy Sets and Systems*, **20**, 1986, 343-353
KAUFMANN, A (1975) *Introduction to the Theory of Fuzzy Subsets* Vol I Academic Press, New York
MATHIEU-NICOT, B (1986) Fuzzy Expected Utility *Fuzzy Sets and Systems* **19**, 163-173
YAGER, R and BASSON, D (1976) Decision-making with Fuzzy Sets *Decision Sciences* **6**, 590–600
ZADEH, L (1965) Fuzzy Sets *Inf and Control* **8**, 338–353
ZADEH, L, FU, K, TANAKA, K and SHIMURA, M, Editors (1975) *Fuzzy Sets and Their Applications to Cognitive and Decision Processes* Academic Press, New York
ZIMMERMANN, H-J (1985) Fuzzy Set Theory and its Applications Kluwer, Boston
ZIMMERMANN, H-J (1987) Fuzzy Sets, Decision Making and Expert Systems Kluwer, Boston

JEAN LEMAIRE
Insurance Department, Wharton School, 3641 Locust Walk,
University of Pennsylvania, Philadelphia, PA 19104, U.S.A.

PSEUDO COMPOUND POISSON DISTRIBUTIONS IN RISK THEORY

By W. Hürlimann

Winterthur, Switzerland

Abstract

Using Laplace transforms and the notion of a pseudo compound Poisson distribution, some risk theoretical results are revisited A well-known theorem by Feller (1968) and Van Harn (1978) on infinitely divisible distributions is generalized. The result may be used for the efficient evaluation of convolutions for some distributions. In the particular arithmetic case, alternate formulae to those recently proposed by De Pril (1985) are derived and shown more adequate in some cases. The individual model of risk theory is shown to be pseudo compound Poisson. It is thus computable using numerical tools from the theory of integral equations in the continuous case, a formula of Panjer type or the Fast Fourier transform in the arithmetic case. In particular our results contain some of De Pril's (1986/89) recursive formulae for the individual life model with one and multiple causes of decrement. As practical illustration of the continuous case we construct a new two-parametric family of claim size density functions whose corresponding compound Poisson distributions are analytical finite sum expressions. Analytical expressions for the finite and infinite time ruin probabilities are also derived.

Keywords

Pseudo compound Poisson, integral equation, infinite divisibility; multiple decrement model, ruin probability.

1. Pseudo compound Poisson distributions

In order to investigate probability density "functions" such as

$$f(x) = \exp(-\lambda)\,\delta(x)+(1-\exp(-\lambda))\,\mu\exp(-\mu x),$$

$$\lambda,\ \mu > 0,\ \delta(x) \text{ the Dirac function,}$$

we need the theory of "generalized functions" or "distributions" in the sense of L. Schwartz (1950/51/65/66). In this paper we refer to the presentation by Doetsch (1976) (English translation is available) To avoid a conflict of terminology between Function Theory and Statistics we use the term generalized function. This is a linear and continuous functional on the space of infinitely differentiable functions on $\mathbb{R}$ with compact support In this paper generalized functions are usually written without argument as $f, g,$. Sometimes and especially in applications we will abuse notation and write $f(x)$

instead of f, e.g. we write $\delta(x)$ for the Dirac function instead of δ. Integrals are always understood in the Lebesgue sense.

Let $\mathcal{L}$ be the space of all locally integrable functions on $[0, \infty)$ (i.e. integrable in every finite subinterval of $[0, \infty)$), and let $\mathcal{D}$ be the space of all generalized functions on $\mathbb{R}$. For $f \in \mathcal{L}$, $s \in \mathbb{C}$, the *Laplace transform* of $f(x)$ is defined to be

$$Lf(s) = \int_0^\infty \exp(-st)\, f(t)\, dt .$$

This mathematical object is extended as follows to an appropriate subspace of $\mathcal{D}$ (see DOETSCH, § 12). Let D^k, $k = 1, 2, \ldots$, be the k-th derivative operator acting on the space $\mathcal{D}$. A generalized function f is said to be of *finite order k* if $f = D^k h(x)$ for a continuous function $h(x)$ defined on $\mathbb{R}$, and k is the smallest integer with this property. For example, the Dirac function

$$\delta = D^2 h(x), \ h(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

is of order 2. Restrict now $\mathcal{D}$ to the subspace $\mathcal{D}_0$ of generalized functions of finite order whose associated continuous functions $h(x)$ satisfy the conditions

$$\begin{aligned} &h(x) = 0 \qquad \text{for} \qquad x < 0, \\ &Lh(s) \text{ converges absolutely for } Re(s) > \sigma, \\ &\sigma \text{ dependent on } h. \end{aligned}$$

For $f = D^k h(x) \in \mathcal{D}_0$, $s \in \mathbb{C}$, the *Laplace transform* is defined to be

$$Lf(s) = s^k Lh(s) \tag{1.1}$$

and is an analytical function for $Re(s) > \sigma$. The space $\mathcal{L}$ is embedded in $\mathcal{D}$ as follows. The generalized function defined by $f \in \mathcal{L}$ is the functional

$$\int_{-\infty}^{\infty} f(x)\, \varphi(x)\, dx, \ \varphi(x) \text{ infinitely differentiable on } \mathbb{R} \text{ with compact support.}$$

A function $f \in \mathcal{L}$ with a Laplace transform in the classical sense has the same Laplace transform in the generalized sense (DOETSCH, Satz 12.2). Moreover the inverse of the Laplace transform is unique up to a zero (generalized) function in both the classical and generalized sense (DOETSCH, Satz 5.1, and p 72) Here the zero function $z(x)$ in $\mathcal{L}$ is a function such that

$$\int_0^t z(x)\, dx = 0, \qquad \text{for all} \qquad t \geq 0 .$$

The convolution operator on $\mathcal{D}_0$ is defined as follows. If $f = D^m h(x)$, $g = D^n k(x)$, then

$$f * g = D^{m+n} (h * k)(x) .$$

The operations on the classical Laplace transform extend to the generalized case Some operations used in this paper are summarized in the next Table

(Generalized) function	*Laplace transform*
f, g	$Lf(s),\ Lg(s)$
$af+bg,\ a, b \in \mathbf{R}$	$aLf(s)+bLg(s)$
$f * g$	$Lf(s)\ Lg(s)$
xf	$-(d/ds)\,Lf(s)$
$\exp(-ax)f,\ a \in \mathbf{R}$	$Lf(s+a)$
$f'(x)$	$sLf(s)-f(0^+)$
$\delta(x)$ (Dirac function)	1

To illustrate the consistency of the Table with definition (1.1) we derive the formula for the Laplace transform of the n-th derivative $f^{(n)}$ of a function $f \in \mathcal{Y}$. From the theory of generalized functions (e.g. DOETSCH, § 14) one knows that

$$D^n f = f^{(n)} + f^{(n-1)}(0^+)\,\delta + \ldots + f(0^+)\,\delta^{(n-1)}$$

Since $L\delta^{(k)}(s) = s^k$ it follows with (1.1) that

$$\begin{aligned} s^n\, Lf(s) &= LD^n f(s) \\ &= Lf^{(n)}(s) + f^{(n-1)}(0^+) + \ldots + f(0^+)\, s^{n-1}, \end{aligned}$$

which provides after rearrangement the desired formula. The differential rule for a generalized function $f \in \mathcal{Y}_0$ looks somewhat different, namely

$$LD^n f(s) = s^n\, Lf(s).$$

From now on our main concern is probabilistic. The set of locally integrable probability density functions $f \in \mathcal{Y}$ is denoted by $\mathcal{Y}P$ The distribution corresponding to $f(x)$ is

$$F(x) = \int_0^x f(t)\, dt.$$

It s well-known that Panjer's recursive formula plays an important role in computational risk theory. For $f \in \mathcal{Y}P$ we are interested in the analogous integral equation

$$\text{(1 2)} \qquad xf(x) = \lambda \int_0^x yh(y)\, f(x-y)\, dy, \qquad \lambda \in \mathbb{R},$$

where $h \in \mathcal{Y}$ is not necessarily positive In applications of risk theory the assumption $0 < F(0) < 1$ is almost always fulfilled. We consider therefore the subset $\mathcal{Y}P_0$ of all functions $f \in \mathcal{Y}P$ with $0 < F(0) < 1$ and for which there is a unique solution $h \in \mathcal{Y}$ with

$$\int_0^\infty h(x)\,dx = 1,$$

such that (1.2) is almost everywhere fulfilled From results by STEUTEL (1970) and VAN HARN (1978) the set $\mathcal{C}\,P_o$ contains all infinitely divisible densities on $(0, \infty)$ (see Corollary 2). It has been shown in the arithmetic case that there are interesting non-infinitely divisible distributions on $\mathbb{N}$ for which the arithmetic version of (1.2) is fulfilled, e.g. the individual model of risk theory with multiple causes of decrement (HÜRLIMANN (1989b)) Are there analogous continuous candidates in $\mathcal{C}\,P_o$ and what is exactly this set? A practical answer is postponed to the end of this Section. From a mathematical point of view, the set $\mathcal{C}\,P_o$, given that it contains non-infinitely divisible distributions, is appealing, since it leads to a natural generalization of the characterization by FELLER (1968) and VAN HARN (1978) of infinitely divisible distributions with non-vanishing zero-probability.

THEOREM 1 Let $f(x)$ be in the class $\mathcal{C}\,P_o$. Then in the space $\mathcal{D}_o$ the following representation holds almost everywhere

$$f(x) = \sum_{k=0}^{\infty} \exp(-\lambda)\,\lambda^k/k!\; h^{*k}(x)$$

where $h^{*o}(x) = \delta(x)$, $\lambda = -\ln\{F(0)\}$ and $h(x)$ is almost everywhere the unique solution of the integral equation

$$xf(x) = \lambda \int_0^x yh(y)\,f(x-y)\,dy. \tag{1.3}$$

PROOF. The integral equation (1 3) can be rewritten as

$$xf(x) = \lambda(f * u)(x) \qquad \text{with} \qquad u(x) = xh(x).$$

Applying the Laplace transform we get

$$(d/ds)\,Lf(s) = \lambda\, Lf(s)\cdot(d/ds)\,(Lh(s))$$

It follows that

$$Lf(s) = c\cdot\exp(\lambda\, Lh(s)).$$

By Laplace inversion in the space $\mathcal{D}_o$ we get almost everywhere

$$f(x) = c\sum_{k=0}^{\infty} \lambda^k/k!\; h^{*k}(x).$$

In this formula we see that $p = f - c\delta \in \mathcal{D}_o$ comes from a function $p \in \mathcal{C}$ By integration

$$F(x) = c + \int_0^x p(t)\,dt,$$

which shows that $c = F(0)$. Put $\lambda = -\ln\{F(0)\}$ to get the result.

The above result suggests the following definition.

DEFINITION. A probability density function $f(x)$ defined on $(0, \infty)$ is said to be of *pseudo compound Poisson* type if $f \in \mathscr{C}P_0$. We call the associated $h(x)$ a *pseudo density*.

INTERPRETATION In risk theory and when it is actually non-negative the function $h(x)$ plays the role of claim size density

The following equivalent formulation of Theorem 1 can be more adequate for practical evaluations In particular it generalizes the result by STRÖTER (1985)

COROLLARY 1. Let $f(x)$ be pseudo compound Poisson with parameter λ and pseudo density $h(x)$ Define $p(x) = f(x) - \exp(-\lambda)\,\delta(x)$.

Then $p(x)$ satisfies the integral equation

$$xp(x) = \lambda \exp(-\lambda)\, xh(x) + \lambda \int_0^x yh(y)\, p(x-y)\, dy \tag{1 4}$$

PROOF. Introduce $f(x) = \exp(-\lambda)\,\delta(x) + p(x)$ in the integral equation (1.3) to obtain immediately (1 4).

In view of its importance both in theory and practice (see e g. STEUTEL (1979)) we recall the definition of infinite divisibility.

DEFINITION. A random variable X, taking values in $\mathbb{R}$, is called *infinitely divisible* if for every $n \in \mathbb{N}$ there exist independent, identically distributed random variables $Y_{1,n}, \ldots, Y_{n,n}$ such that the following equality in distribution is valid·

$$X \stackrel{d}{=} Y_{1,n} + \ldots + Y_{n,n}.$$

Equivalently $P(z)^{1/n} = E[z^X]^{1/n}$, $Lf(s)^{1/n} = E[\exp\,(-sX)]^{1/n}$ or $\varphi(t)^{1/n} = E[\exp(itX)]^{1/n}$ is respectively a probability generating function, a Laplace transform or a characteristic function for every $n \in \mathbb{N}$ The associated probability density and distribution are also called infinitely divisible

The special case of Theorem 1 for infinitely divisible distributions on $[0, \infty)$ has been identified in other forms by STEUTEL (1970) and VAN HARN (1978) in the general and KATTI (1967) and FELLER (1968) in the arithmetic case

COROLLARY 2. Let X be a random variable defined on $[0, \infty)$ with locally integrable density $f(x)$ such that $0 < F(0) < 1$. Then the following conditions are equivalent·

(a) X is infinitely divisible;
(b) X is compound Poisson with parameter λ and jump density $h(x)$ and $f(x)$ is solution of the integral equation (1.3);
(c) The solution $h(x)$ of the integral equation (1.3) is positive

PROOF. In the arithmetic case the equivalence of (a) and (c) has been shown by KATTI (1967) (other proof by STEUTEL (1970, p. 83)) The equivalence of (a) and (b) was shown by FELLER (1968, vol. 1, 3rd edition, p. 290) (other proof by GERBER and VALDERRAMA OSPINA (1987)). In the continuous case the equivalence of (a) and (b) is due to VAN HARN (1978, theorem 1 6.6) for the compound Poisson representation and STEUTEL (1970) (see also VAN HARN, Corollary 1 6.3) for the integral equation representation. The equivalence of (b) and (c) follows from Theorem 1.

Next we display a subclass of $\mathcal{D}P_0$ which is big enough for our applications. In particular we will show by construction in Section 4 that the class $\mathcal{D}P_0$ contains more functions than the infinitely divisible ones.

THEOREM 2. Let $\mathcal{D}P'$ be the subclass of $\mathcal{D}P$ consisting of functions $f(x)$ which satisfy the following conditions:

(i) $0 < F(0) < 1$.
(ii) The associated generalized function $f - F(0)\,\delta \in \mathcal{D}'_0$ comes from a continuous function $f(x) - F(0)\,\delta(x)$ defined on $[0, \infty)$.

Then $\mathcal{D}P'$ is contained in $\mathcal{D}P_0$.

PROOF Let $f \in \mathcal{D}P'$. The function $p(x) = f(x) - F(0)\,\delta(x)$ is by assumption continuous on $[0, \infty)$ Consider the Volterra integral equation of the second kind

$$a(x) = \lambda^{-1} \exp(\lambda)\, x p(x) - \exp(\lambda) \int_0^x a(y)\, p(x-y)\, dy, \quad \lambda = -\ln\{F(0)\}.$$

Since $p(x-y)$ and $xp(x)$ are continous on $\{0 \le x \le a, 0 \le y \le x\}$ respectively $\{0 \le x \le a\}$, this equation can be solved uniquely for $a(x)$ (see e g JERRI (1985), p. 194 and p 201) Set $h(x) = a(x)/x$. After algebraic manipulation one sees that $h(x)$ is the unique solution of the integral equation (1.4). Since $f(x) = F(0)\,\delta(x) + p(x)$, one checks easily that $h(x)$ is also the unique solution of the integral equation (1.3). Provided that

$$\int_0^\infty h(x)\, dx = 1,$$

we have shown that $f \in \mathcal{D}P_0$. This point is proved as follows. Since $h(x)$ is solution of (1 3) one shows that

$$\int_0^\infty h(x)\,dx = c < \infty$$

Then $\bar{h}(x) = h(x)/c$ is the unique solution of the integral equation

$$xf(x) = \lambda c \int_0^\infty y\bar{h}(y)\, f(x-y)\,dy.$$

Since $\int_0^\infty \bar{h}(x)\,dx = 1$ one has $f \in\!\!/\; P_o$. But from Theorem 1 one has then

$$\lambda c = -\ln\{F(0)\}.$$

By definition of λ above one has indeed $c = 1$.

REMARKS

(1) In Theorem 1 and Corollary 2 the condition $F(0) > 0$ is necessary. The infinitely divisible exponential density $f(x) = \mu\exp(-\mu x)$ leads to the solution $h(x) = \exp(-\mu x)/x$, but $\int_0^\infty h(x)\,dx = \infty$. This density is not compound Poisson, but the weak limit of the compound Poisson densities $f_\lambda(x) = \exp(-\lambda)\,\delta(x) + (1-\exp(-\lambda))\,\mu\exp(-\mu x)$ as $\lambda \to \infty$, with claim size densities $h_\lambda(x) = \exp(-\mu x)\,(1-\exp(-ax))/\lambda x$, $a = (\exp(\lambda)-1)\,\mu$. This result will be derived in Section 4. In general $p(x)$ with $P(0) = 0$ is infinitely divisible if and only if $f_\lambda(x) = \exp(-\lambda)\,\delta(x) + (1-\exp(-\lambda))\,p(x)$ is infinitely divisible with $F(0) = \exp(-\lambda)$ and $p(x)$ is the weak limit of the f_λ's as $\lambda \to \infty$. (FELLER (1968), vol 2, 2nd edition, p. 303).

(2) In the arithmetic case the integral equation (1.3) is to be replaced by the well-known Panjer recursive formula

$$kp(k) = \lambda \sum_{s=1}^{k} sh(s)\,p(k-s) \qquad (1.5)$$

An independent and more elementary proof of the results in this mathematically simpler case in presented in HÜRLIMANN (1989a, 1989b). Observe that Laplace transforms are to be replaced by the geometric transform (= probability generating function in case of arithmetic distributions, see GIFFIN (1975) for fundamentals)

(3) Methods to solve integral equations can be found in all parts of Applied Mathematics. Transform theory (see WIDDER (1971)), especially Laplace transforms, is a powerful tool to get closed analytical results An illustration is given in Section 4. Numerical methods were extensively studied by BAKER (1977) and more recently equation (1 4) has been solved in the insurance context by STRÖTER (1985). It is worthwile to mention that the Laplace

transform approach simplifies the derivation of Theorem 1 1. of the latter author, which uses the method of successive approximation.

(4) Theorem 1 can be interpreted as a duality assertion. There is a duality between integrable densities on $[0, \infty)$ and pseudo densities, where the pseudo compound Poisson representation realizes this duality The subclass of infinitely divisible densities is just dual to the ordinary densities.

(5) Theorem 1 suggests many (also difficult) applications. It can be useful for the computational evaluation of convolutions (see next Section), as well as for the study of other properties of exact sampling distributions. A statistical application is given in HÜRLIMANN (1989a).

(6) With more technical refinements it should be possible to extend the results to arbitrary one-sided unbounded intervals $[a, \infty)$, $a > -\infty$, (see VAN HARN (1978) for the case of infinitely divisible distributions). It would be of great interest to generalize Theorem 1, if possible, to the whole real line and especially obtain a single characterizing functional equation valid on $\mathbb{R}$. Unfortunately, even for infinitely divisible distributions, the latter requirement is still an open problem, as reported by VAN HARN (1978), p. 189.

2. CONVOLUTIONS OF DISTRIBUTIONS

Let $X_1, X_2, \ldots, X_n$ be n mutually independent random variables on $[0, \infty)$ with a common integrable density $f(x)$ such that $0 < F(0) < 1$. In probability and statistical theory one is interested in the exact sample distribution of the mean. It is a straightforward rescaling of the distribution of the sum

$$X = X_1 + \ldots + X_n$$

whose density is given by the n-fold convolution

$$\bar{f}(x) = f^{*n}(x).$$

The evaluation of this function uses the recursive formula

$$f^{*(k+1)}(x) = \int_0^x f(y) f^{*k}(x-y)\, dy,$$

which is very time-consuming for large values of n, especially when $f(x)$ is not a simple function

Using Theorem 1 and the various methods for solving integral equations, an alternative general approach to this problem follows immediately. In the following we will often use $g(x) = \lambda h(x)$ instead of $h(x)$

COROLLARY 3. Let the X_i be defined on $[0, \infty)$ with $0 < F(0) < 1$. Assume $f \in \mathscr{P}P_o$. Let $g(x)$ be the solution of the integral equation

(2.1) $$xf(x) = \int_0^x yg(y)\, f(x-y)\, dy$$

Then the n-fold convolution $\bar{f}(x)$ is solution of the integral equation

(2.2) $$x\bar{f}(x) = n \int_0^x yg(y)\, \bar{f}(x-y)\, dy$$

PROOF. In the proof of theorem 1 we have seen that

$$Lf(s) = F(0) \exp(Lg(s)),$$

and thus

$$L\bar{f}(s) = F(0)^n \exp(nLg(s)).$$

Therefore $\bar{f}(x)$ is pseudo compound Poisson with parameter $n\lambda$ and pseudo density $g(x)/\lambda$. The affirmation follows from Theorem 1.

Let us have a look to the special arithmetic case. The n-fold convolution $\bar{p}(x) = p^{*n}(x)$ can be evaluated using the recursive Panjer formula

(2.3) $$\begin{cases} \bar{p}(0) = p(0)^n \\ k\bar{p}(k) = n \sum_{s=1}^{k} sg(s)\, \bar{p}(k-s) \end{cases}$$

where $g(s)$ is itself computed recursively by

(2.4) $$sg(s)\, p(0) = sp(s) - \sum_{i=1}^{s-1} ig(i)\, p(s-i)$$

At first sight it might appear that this two-stage nested recursive algorithm is computationally less efficient than the recursive formula proposed by DE PRIL (1985), Theorem 1

(2.5) $$\begin{cases} \bar{p}(0) = p(0)^n \\ k\bar{p}(k)\, p(0) = \sum_{s=1}^{k} [(n+1)s-k]\, p(s)\, \bar{p}(k-s) \end{cases}$$

In some cases it might be that only $g(k)$ is known and $p(k)$ must be computed recursively using Panjer's formula (1.5). Then the formula (2.3) is simpler and more direct than formula (2.5)

EXAMPLES. The choice

(2.6) $$g(k) = \frac{p \cdot \Gamma(a+k-1)\, c^{k-1}}{\Gamma(a)\, k!\; (1+c)^{a+k-1}}, \quad k = 1, 2, \ldots, p > 0,\ c > 0,\ a \geq 0$$

leads to Hoffmann/Thyrion's family proposed as claim number distribution by KESTEMONT and PARIS (1985/87). A similar choice would be the ETNB distribution

(2.7) $$g(k) = \frac{\Gamma(k+a)\ \beta^k}{\Gamma(a)k!\ [(1-\beta)^{-a}-1]}, \quad k = 1, 2, \ldots, \; -1 < a < 0, 0 < \beta \le 1,$$

studied as probability density (however) by WILLMOT (1988). In these examples it is more direct to apply formula (2 3) to compute exact n-fold convolutions than to use De Pril's formula (2.5).

3. THE INDIVIDUAL MODEL OF RISK THEORY

Consider n mutually independent random variables $X_1, X_2, \ldots, X_n$, not necessarily identically distributed as in Section 2 Suppose each X_i has a range contained in the interval $[0, \infty)$, which may be arithmetic or not. In risk theory the sum

$$X = X_1 + X_2 + \ldots + X_n,$$

called individual model, can be interpreted as the aggregate claims in a finite period on a portfolio of n independent contracts. Let $F(x) = \Pr(X \le x)$, $F_i(x) = \Pr(X_i \le x)$, $i = 1, 2, \ldots, n$, and assume that $0 < F_i(0) < 1$ for all i.

THEOREM 2. Assume the probability densities $f_i \in \mathscr{P}_0$, $i = 1, \ldots, n$ Then the individual model of risk theory is pseudo compound Poisson with parameter

(3.1) $$\lambda = -\ln\{F(0)\} = -\sum_{i=1}^{n} \ln\{F_i(0)\},$$

and pseudo density

(3 2) $$h(x) = \left(\sum_{i=1}^{n} g_i(x)\right)/\lambda,$$

where each $g_i(x)$ is unique solution of the integral equation

(3.3) $$xf_i(x) = \int_0^x y g_i(y)\, f_i(x-y)\, dy$$

PROOF Clearly $f = f_1 * f_2 * \ldots * f_n$. In the proof of Theorem 1 we have seen that

$$Lf_i(s) = F_i(0)\ \exp(Lg_i(s)), \qquad i = 1, 2, \ldots, n$$

It follows that

$$Lf(s) = \sum_{i=1}^{n} Lf_i(s) = \prod_{i=1}^{n} F_i(0)\ \exp\left(\sum_{i=1}^{n} Lg_i(s)\right)$$

Taking inverse Laplace transforms in the space $\mathscr{V}_0$ the result follows immediately

For simplicity restrict the following discussion to the arithmetic case First of all formulae for $g_i(x)$ must be obtained, or the $g_i(x)$ must be computed by other means, using for example Panjer's recursive formula (3.3). Then the probability density function of the individual model can be computed using Panjer's recursion, valid in the generalized case

$$(3.4)\qquad f(x) = \begin{cases} \prod_{i=1}^{n} f_i(0), & x = 0, \\ (-\ln\{f(0)\}/x) \sum_{y=1}^{x} y h(y)\, f(x-y), & x > 0. \end{cases}$$

Compared to the collective model of risk theory the extra cost for preparing $h(x)$ may be substantial since many values of $g_i(x)$, $i = 1, 2, \ldots$, are involved in the computation. A sound procedure would be to approximate the pseudo density, as suggested by DE PRIL (1987/89) (see Example 1 below), by a more tractable function $h^*(x)$ and compute the approximate density

$$(3.5)\qquad f^*(x) = \begin{cases} \prod_{i=1}^{n} f_i(0), & x = 0, \\ (-\ln\{f(0)\}/x) \sum_{y=1}^{x} y h^*(y)\, f^*(x-y), & x > 0 \end{cases}$$

Another possibility to reduce the computational effort is to apply the Fast Fourier Transform, inverting the Fourier transform of the pseudo compound Poisson representation according to the formula

$$\bar{f} = \{f(0)/n\}\, \mathrm{FFT}^{-}(\exp(\mathrm{FFT}^{+}(\bar{g})))$$

Here FFT^{+}, $\{1/n\}\,\mathrm{FFT}^{-}$ denote Fast Fourier Transform, respectively the inverse transform, and n is the size of the vectors $\bar{f}$, $\bar{g}$ associated to the functions $f(x)$, $g(x)$. Since one has to take into accout a relatively long support of $h(x)$, the FFT-method has been shown superioi to Panjer's recursion in many cases (cf BÜHLMANN (1984)), and the error bound in the distribution as well as in associated stop-loss premiums are controllable (BUHLMANN (1984), HURLIMANN (1986))

EXAMPLE 1. The simplest individual life model has been considered by DE PRIL (1986/87). Let n_{ij} be the number of policies with amount at risk i and mortality rate q_j, $i = 1, \ldots, a$, $j = 1, \ldots, b$ Let $p_j = 1 - q_j$ the corresponding survival probabilities, $n_j = \sum_{i=1}^{a} n_{ij}$ the number of policies with mortality rate q_j, $n = \sum_{j=1}^{b} n_j$ the total number of policies, and $m = \sum_{i=1}^{a} \sum_{j=1}^{b} i \cdot n_{ij}$ the maximum possible amount of aggregate claims. Furthermore let X_{ij} be the random variable representing the claim produced by a policy with amount at risk i and mortality rate q_j. Its probability density function is given by

$$f_{ij}(x) = \begin{cases} p_j, & x = 0 \\ q_j, & x = i \\ 0, & \text{else} \end{cases} \tag{3.6}$$

Following the device given by the arithmetic version of Theorem 2 we search for unique functions $g_{ij}(x)$ such that

$$x f_{ij}(x) = \sum_{y=1}^{x} y g_{ij}(y)\, f_{ij}(x-y)$$

In the lemma below they are shown to be

$$g_{ij}(x) = \begin{cases} (-1)^{k-1}/k \cdot (q_j/p_j)^k, & x = ik, \quad k = 1, 2, \ldots \\ 0, & \text{else} \end{cases} \tag{3.7}$$

It follows that this individual model is pseudo compound Poisson with parameter

$$\lambda = -\sum_{j=1}^{b} n_j \ln(p_j) = -\ln\{f(0)\}$$

and pseudo density

$$h(x) = 1/\lambda \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} g_{ij}(x).$$

Insert these formulae in (3.4). Then one has

$$f(0) = \prod_{j=1}^{b} (p_j)^n$$

For $x > 0$ one obtains with $y = ik$ ·

$$x f(x) = \sum_{i=1}^{\min(a,x)} \sum_{k=1}^{[x/i]} A(i,k)\, f(x-ik), \qquad x = 1, 2, \ldots, m \tag{3.8}$$

with

$$A(i,k) = (-1)^{k+1} i \sum_{j=1}^{b} n_{ij} (q_j/p_j)^k.$$

This has been derived differently by De Pril (1986). For computational reasons Reimers (1988) has proposed to reverse the order of summation:

(3.9) $$xf(x) = \sum_{k=1}^{x} \sum_{i=1}^{\min(a,[x/k])} A(i,k)\, f(x-ik)$$

To save computer time it is advisable to truncate the first summation taking only 4-5 terms as proposed by De Pril and Vandenbroek (1987). An analysis of the magnitude of error involved in this approximation step is given by De Pril (1988).

Lemma The Panjer recurrence relation equations

$$xf(x) = \sum_{y=1}^{x} yg(y)\, f(x-y)$$

where

$$f(x) = \begin{cases} p, & x = 0 \\ q, & x = i, \qquad 0 < q < 1, \qquad p+q = 1, \\ 0, & \text{else} \end{cases}$$

have the unique solution

$$g(x) = \begin{cases} (-1)^{k-1}/k \cdot (q/p)^k, & x = ik, \quad k = 1, 2, \ldots \\ 0, & \text{else} \end{cases}$$

Proof One uses induction. For this rewrite the recurrence equations in form (2.4):

$$xg(x)\, f(0) = xf(x) - \sum_{y=1}^{x-1} yg(y)\, f(x-y).$$

For $x = 1, \ldots, i-1$ one obtains $g(x) = 0$ For $x = i$ the equation reads

$$ig(i)p = iq.$$

Hence one has $g(i) = q/p$. Let now $x > i$ and assume the formula for $g(y)$ correct for all $y < x$. If $x = ik$ is a multiple of i, then the right-hand side of the equation gives a contribution only for $x-y = i$, that is $y = (k-1)i$. The equation reads

$$ikg(x)p = -(k-1)\, ig((k-1)i)q$$

and the correct value of $g(x)$ is checked by induction assumption. When x is not a multiple of i the right-hand side vanishes and hence $g(x) = 0$.

Example 2. Consider the individual life multiple decrement model which has applications in pension theory for example (see Bowers et al. (1986)). Let m be the number of causes of decrement and let the vector $\underline{s} = (s_1, \ldots, s_m)$ represent amounts at risk, s_j being a sum at risk due to cause j. The vector $\underline{s}$ is assumed to take values in a finite set $A \subset \mathbb{Z}^m$. Let $n_{\underline{s}k}$ be the number of policies with risk sum structure $\underline{s}$ and probabilities of decrement $q_k^{(j)}$ due to cause, j, $j = 1, \ldots, m$, $k = 1, \ldots, b$ Let $p_k^{(\tau)} = 1 - \sum_{j=1}^{m} q_k^{(j)}$ be the survival probability due to all causes of decrement Denote by $n_k = \sum_{\underline{s} \in A} n_{\underline{s}k}$ the number of policies with survival probability $p_k^{(\tau)}$ and by $n = \sum_{k=1}^{b} n_k$ the total number of policies The maximum possible amount of aggregate claims is denoted by M and is equal to

$$M = \sum_{k=1}^{b} \sum_{\underline{s} \in A} \max_{1 \leq j \leq m} (s_j)\, n_{\underline{s}k}.$$

Moreover let the random variable $X_{\underline{s}k}$ represent the claim produced by a policy with risk sum structure $\underline{s}$ and probabilities of decrement $q_k^{(j)}$, $j = 1, \ldots, m$, $k = 1, \ldots, b$ Its probability density function, denoted by $f_{\underline{s}k}(x)$, is given by

$$(3.10) \qquad f_{\underline{s}k}(x) = \begin{cases} p_k^{(\tau)}, & x = 0 \\ q_k^{(j)}, & x = s_j, \qquad j = 1, \ldots, m, \\ 0, & \text{else} \end{cases}$$

Evaluate now the probability density function of aggregate claims using Panjer's recursive formula (3.4). We have clearly

$$f(0) = \prod_{k=1}^{b} (p_k^{(\tau)})^{n_k}.$$

For $x > 0$ it is necessary to evaluate first in a recursive manner the functions $g_{\underline{s}k}(x)$ such that

$$(3.11) \qquad x f_{\underline{s}k}(x) = \sum_{y=1}^{x} y g_{\underline{s}k}(y)\, f_{\underline{s}k}(x-y), \qquad \underline{s} \in A, \qquad k = 1, \ldots, b.$$

Then

$$(3.12) \qquad h(x) = 1/\lambda \sum_{\underline{s} \in A} \prod_{k=1}^{m} n_{\underline{s}k}\, g_{\underline{s}k}(x),$$

$$\lambda = -\ln\{f(0)\},$$

is introduced in the recursive formula (3.4). It is important to note that the proposed algorithm requires a two-stage nested recursive computation. Up to the maximum possible amount of aggregate claims M prepare for each $y = 1, 2, \ldots, M$ the finite number of elements $g_{\underline{s}k}(y)$ recursively solving (3.11) such that

$$g_{\underline{s}k}(y) = \left[y f_{\underline{s}k}(y) - \sum_{z=1}^{y-1} z g_{\underline{s}k}(z)\, f_{\underline{s}k}(y-z) \right] / (y p_k^{(\tau)}) \tag{3.13}$$

Then apply Panjer's recursive formula (3.4) computing $h(y)$ using formula (3.12). As many of the values $f_{\underline{s}k}(y)$ indeed vanish the summation in (3.13) extends over at most m terms. To illustrate consider the double-decrement model with $m = 2$, for example death and withdrawal or death and disability as causes of decrement. Use for brevity the notation $\underline{s} = (i, j)$ with $A = \{1 \leq i, j \leq a\}$. Assuming $i < j$ (the other cases $i = j$ and $i > j$ are similar) the elements $g_{\underline{s}k}(x)$ are computed more efficiently by the recursive formulae

$$q_{\underline{s}k}(x) = \begin{cases} 0, \quad \text{if} \quad x \in \{1, \ldots, i-1\} \quad \text{or} \\ x \in \{i+1, \ldots, j-1 \mid x \text{ not multiple of } i\} \\ (-1)^{r-1}/r \cdot (q_k^{(1)}/p_k^{(\tau)})^r, \quad \text{if} \quad x = ri \neq j, \\ r \in \{1, \ldots, [j/i]\} \\ g_k^{(2)}/p_k^{(\tau)}, \quad \text{if} \quad x = j \text{ is not multiple of } i \\ q_k^{(2)}/p_k^{(\tau)} + (-1)^{r-1}/r \cdot (q_k^{(1)}/p_k^{(\tau)})^r, \quad \text{if} \\ x = j = ri \quad \text{for} \quad r \in \mathbb{N}, \\ -\,[(x-j)\, g_{\underline{s}k}(x-j)\, q_k^{(2)} + \\ +\,(x-i)\, g_{\underline{s}k}(x-i)\, q_k^{(1)}]/(x p_k^{(\tau)}), \quad \text{if} \quad x > j \end{cases} \tag{3.14}$$

An alternative derivation and additional formulae concerning the individual model of risk theory can be found in DE PRIL (1989).

4 PARAMETRIC AGGREGATE CLAIMS MODELS

It is well-known that the compound Poisson gamma and the compound negative binomial exponential distributions can be expressed as analytical series, the latter one as a finite sum. Other cases are less well-known. For many practical purposes it is most desirable to have tractable parametric functions modeling aggregate claims. The classical approach to this problem uses asymptotic approximate formulae as Normal, Normal-Power, Wilson-Hilferty, three-parameter gamma, Haldane, Esscher transforms and others. These approximations are attached with approximation errors which are usually difficult to control. Furthermore the structure of the claim size density has been lost in these models. Since it is often necessary to study claims frequency and claim size separately, parametric aggregate claims models with explicit structure of claim number and claim size distribution are of interest. This can be achieved solving analytically integral equations of the form (1.4). The method is illustrated at a simple new case, namely a modified two parameter gamma aggregate claims model.

Let $f(x)$ be an aggregate claims density such that $0 < F(0) = \exp(-\lambda) < 1$. This assumption is in particular fulfilled for a Poisson claim number model with parameter λ and when there are no claims of amount ≤ 0. More generally this can be assumed for infinitely divisible aggregate claims distributions defined on $[0, \infty)$ (see Corollary 1). Rewrite the density as

(4.1) $$f(x) = \exp(-\lambda)\,\delta(x) + g(x)$$

The derivative $(d/ds)\,Lf(s)$ of a Laplace transform is denoted for short by $L'f(s)$ Solving the integral equation (1 4) is equivalent to solving a differential equation in the Laplace space and taking inverse Laplace transforms The differential equation reads

(4 2) $$L'g(s) = \lambda Lg(s)\,L'h(s) + \lambda \exp(-\lambda)\,L'h(s)$$

Given the function $h(x)$ its general solution is

(4.3) $$Lg(s) = c \cdot \exp(\lambda Lh(s)) - \exp(-\lambda).$$

where c is a constant. We have gained nothing since this is equivalent to the pseudo compound Poisson representation and is difficult to handle analytically. However specifying the function $g(x)$ it might be easier to find $h(x)$ according to the formula

(4.4) $$L'h(s) = \exp(\lambda)\,L'g(s)/[\lambda(1+\exp(\lambda)\,Lg(s))]$$

For the modified two-parameter gamma aggregate claims model, the task is to find the pseudo density $h(x)$ which corresponds to

(4.5) $$g(x) = (1-\exp(-\lambda))\,\mu^a x^{a-1} \exp(-\mu x)/\Gamma(a), \qquad a \geq 1, \qquad \mu > 0$$

Setting $\omega = 1-\exp(-\lambda)$ one gets

(4.6) $$Lg(s) = \omega(1+s/\mu)^{-a}, \qquad L'g(s) = -(a\omega/\mu)\ (1+s/\mu)^{-a-1}$$

After straightforward calculation it follows that

(4.7) $$L'h(s) = -a\alpha^a/[\lambda(s+\mu)\ ((s+\mu)^a + \alpha^a)],$$

where α is the positive a-th root defined by

(4.8) $$\alpha^a = (\exp(\lambda)-1)\,\mu^a$$

Inverse Laplace transformation yields

(4.9) $$h(x) = \exp(-\mu x)/\lambda x \int_0^x L^{-1}[a\alpha^a/(\alpha^a+s^a)](y)\,dy$$

We show now that for integer values $a = n = 1, 2, 3,$. the function $h(x)$ has a *finite closed form.* Using properties of the Laplace transform it suffices to invert the functions

(4.10) $$L'\bar{h}(s) = -1/[s(1+s^n)] = s^{n-1}/(1+s^n) - 1/s, \qquad n = 1, 2, \ldots$$

Set $\bar{h}(x) = \bar{h}_1(x) + \bar{h}_2(x)$ with $L'\bar{h}_1(s) = -1/s$, $L'\bar{h}_2(s) = s^{n-1}/(1+s^n)$. It follows that $\bar{h}_1(x) = 1/x$, $x > 0$, and $\bar{h}_2(x) = -(1/x) \cdot L^{-1}[s^{n-1}/(1+s^n)](x)$, $x > 0$. To find the latter inverse Laplace transform expand the rational function as a partial fraction (e.g. DOETSCH (1976), p. 89):

(4.11) $$s^{n-1}/(1+s^n) = 1/n \sum_{k=0}^{n-1} 1/(s-\exp(i(2k+1)\pi/n))$$

and re-group the complex conjugate terms. As n is odd or not one obtains two different formulae summarized as follows:

(4.12) $$s^{n-1}/(1+s^n) = (1/n)\Big[(1-(-1)^n)/[2(1+s)] + \sum_{k=0}^{[n/2]-1} 2(s-a_{k,n})/(s^2-2a_{k,n}s+1)\Big]$$

where $a_{k,n} = \cos[(2k+1)\pi/n]$. For later use set $\beta_{k,n} = |\sin[(2k+1)\pi/n]|$. From a table of Laplace transforms (e.g. DOETSCH (1976)) one has

$$L^{-1}[1/(s^2-2as+1)](x) = (1/\beta)\exp(ax)\sin(\beta x).$$

It follows that

(4.13) $$L^{-1}[(2s-2a)/(s^2-2as+1)](x) = 2\exp(ax)\cos(\beta x)$$

whenever $a^2+\beta^2 = 1$. Using these results one gets after some algebraic manipulation the pseudo density in form of a finite sum:

(4.14) $$h(x) = (\exp(-\mu x)/\lambda x)\Big[n-(1-(-1)^n)\exp(-ax)/2 - \sum_{k=0}^{[n/2]-1} 2\exp(a_{k,n}ax)\cos(\beta_{k,n}ax)\Big]$$

with $a = (\exp(\lambda)-1)^{1/n}\mu$. In particular for lower dimensions one has the pseudo densities

(4.15)
$$\begin{aligned} n = 1:\ & h(x) = \exp(-\mu x)(1-\exp(-ax))/(\lambda x), \\ & a = \mu(\exp(\lambda)-1), \\ n = 2:\ & h(x) = 2\exp(-\mu x)(1-\cos(ax))/(\lambda x), \\ & a = \mu\sqrt{\exp(\lambda)-1}, \\ n = 3:\ & h(x) = [\exp(-\mu x)/(\lambda x)][3-\exp(-ax) - \\ & \qquad - 2\exp(ax/2)\cos\{(\sqrt{3}/2)ax\}], \\ & a = \mu\sqrt[3]{\exp(\lambda)-1} \end{aligned}$$

We apply now Corollary 2. For $n = 1, 2$ we have $h(x) > 0$ and the corresponding model (4.1) is infinitely divisible and thus compound Poisson. For $n = 3$ one may have $h(x) < 0$. Hence (4.1) is not infinitely divisible and thus only pseudo compound Poisson. In particular we have shown that the classe $\mathscr{CP}$ is bigger than the class of infinitely divisible probability density functions defined on $(0, \infty)$. As known to the author the present model $n = 1$

is among the few examples of compound Poisson models allowing *finite analytical sum expressions* for the main risk theoretical quantities of interest. In particular it is comparable to the Poisson exponential aggregate claims model concerning mathematical simplicity.

Furthermore analytical expressions for the finite and infinite time ruin probabilities can be derived. We have computed the simple case $n = 1$ (details of calculation in appendix). Assume a stationary evolution of the portfolio. In this context $P = (1+\theta)\lambda m$ represents the premiums received continuously per unit of time, with θ the security loading, m the expected claim size, and λ measures the expected number of claims per unit of time Then the probability of ruin $\psi(x, t)$ before time t given the initial reserves x is

(4.16) $$\psi(0, t) = 1/(1+\theta) - (1-\exp(-\lambda t)) \exp(-\mu Pt)/(\mu Pt),$$

and for $x > 0$,

(4.17) $$\begin{aligned}\psi(x, t) = {} & (1-\exp(-\lambda t)) \exp(-\mu(x+Pt)) + \\ & + \theta/(1+\theta)\cdot\exp(-\mu x)\cdot[\lambda/(\lambda+P\mu) - \\ & - \exp(-\mu Pt)\cdot\{1-\exp(-\lambda t)\cdot P\mu/(\lambda+P\mu)\}] + \\ & + \exp(-\mu(x+Pt)) \sum_{k=2}^{\infty} (-\lambda t)^k/k! \sum_{j=1}^{k-1} 1/j .\end{aligned}$$

Taking limits as $t \to \infty$ it follows that the infinite time ruin probabilities are

(4.18) $$\begin{aligned}\psi(0) &= 1/(1+\theta), \\ \psi(x) &= \theta/(1+\theta)\cdot\exp(-\mu x)\cdot\lambda/(\lambda+P\mu), \qquad x > 0 .\end{aligned}$$

The obtained results will practically be more useful if one fits the claim size density by a linear combination of densities as follows:

(4 19) $$\begin{aligned} h(x) &= \sum_{i=1}^{r} c_i h_i(x), \qquad c_1 + \ldots + c_n = 1, \\ h_i(x) &= \exp(-\mu_i x)\cdot\{1-\exp(-a_i x)\}/\lambda x, \\ a_i &= (\exp(\lambda)-1)\mu_i .\end{aligned}$$

From the proof of Theorem 1 we know that the aggregate claims density $f(x, t)$ up to time t satisfies the Laplace representation

$$(Lf)(s) = \exp(-\lambda t) \exp(\lambda t Lh(s)) = \prod_{i=1}^{r} \exp(-\lambda c_i t)\cdot\exp(\lambda c_i t Lh_i(s)).$$

Define $f_i(x, t)$ as solution of the Laplace equation

$$(Lf_i)(s) = \exp(-\lambda c_i t)\cdot\exp(\lambda c_i t Lh_i(s)).$$

As we have shown, one obtains by inversion

(4.20) $$f_i(x, t) = \exp(-\lambda c_i)\,\delta(x) + (1-\exp(-\lambda c_i t))\cdot\mu_i \exp(-\mu_i x).$$

The direct calculation of the convolutions

$$f(x, t) = f_1(x, t) * \dots * f_r(x, t)$$

yields the formula (use induction):

(4.21) $$f(x, t) = \exp(-\lambda t)\,\delta(x) + \sum_{i=1}^{r} (1-\exp(-\lambda c_i t)) \times$$
$$\times \left[\prod_{j \neq i} (\mu_j - \mu_i \exp(-\lambda c_j t))/(\mu_j - \mu_i)\right] \times$$
$$\times\ \mu_i \exp(-\mu_i x)$$

In this model the net stop-loss premiums to the priority M can be expressed as finite analytical sums, namely

(4.22) $$SL(F, M) = \int_M^\infty (x-M)\, f(x, t)\, dx = \sum_{i=1}^{r} (1-\exp(-\lambda c_i t)) \times$$
$$\times \left[\prod_{j \neq i} (\mu_j - \mu_i \exp(-\lambda c_j t))/(\mu_j - \mu_i)\right] \times$$
$$\times\ \exp(-\mu_i M)/\mu_i$$

Analytical formulae for the finite and infinite time ruin probabilities can also be derived

APPENDIX: CALCULATION OF RUIN PROBABILITIES

Assume an aggregate claims distribution function up to time t of the form

$$F(x, t) = 1-(1-\exp(-\lambda t))\cdot\exp(-\mu x).$$

Then the probability of survival to time t, denoted by $U(x, t) = 1-\psi(x, t)$, can be calculated using Seal's formulae (e.g GERBER (1979)).

$$U(0, t) = \theta/(1+\theta) + (1/Pt)\int_{Pt}^\infty (1-F(z, t))\, dz$$

$$U(x, t) = F(x+Pt, t) - P\int_0^t U(0, t-w)\, f(x+Pw, w)\, dw$$

One obtains

$$U(0, t) = \theta/(1+\theta) + (1-\exp(-\lambda t))\cdot\exp(-\mu Pt)/(\mu Pt).$$

Further calculate

$$U(x, t) = 1-(1-\exp(-\lambda t))\cdot\exp(-\mu(x+Pt)) -$$
$$- P\int_0^t [\theta/(1+\theta)+(1-\exp(-\lambda(t-w)) \times$$
$$\times \exp(-\mu P(t-w))/(\mu P(t-w))] \times$$
$$\times [\exp(-\lambda w)\,\delta(x+Pw)+(1-\exp(-\lambda w)) \times$$
$$\times \exp(-\mu(x+Pw))]\,dw.$$

Since $x+Pw > 0$ for $w \in (0, t)$ the term in $\delta(x+Pw)$ does not contribute to the integral. For clearness write

$$U(x, t) = 1-(1-\exp(-\lambda t))\cdot\exp(-\mu(x+Pt))+I_1+I_2,$$

with

$$I_1 = -P\int_0^t \theta/(1+\theta)\cdot(1-\exp(-\lambda w)) \times$$
$$\times\ \mu\exp(-\mu(x+Pw))\,dw,$$

$$I_2 = -P\int_0^t (1-\exp(-\lambda(t-w)))\cdot(1-\exp(-\lambda w)) \times$$
$$\times\ \exp(-\mu(x+Pt))/(P(t-w))\cdot dw$$

The evaluation of the first integral gives

$$I_1 = \theta/(1+\theta)\cdot\exp(-\mu x)\cdot\left[-P\mu\int_0^t \exp(-\mu Pw)\,dw + \right.$$
$$\left. + P\mu\int_0^t \exp(-(\lambda+P\mu)w)\,dw\right]$$
$$= \theta/(1+\theta)\ \exp(-\mu x)\ [\exp(-\mu Pt) -$$
$$-\ 1+P\mu/(\lambda+P\mu)\cdot(1-\exp(-(\lambda+P\mu)t))]$$
$$= \theta/(1+\theta)\ \exp(-\mu x)\cdot[\exp(-\mu Pt) \times$$
$$\times \{1-\exp(-\lambda t)\cdot P\mu/(\lambda+P\mu)\}-\lambda/(\lambda+P\mu)]$$

To evaluate the second integral expand the first exponential function in a Taylor series to get

$$I_2 = -\exp(-\mu(x+Pt))\sum_{k=0}^{\infty}(-1)^k\lambda^{k+1}/(k+1)!$$
$$\int_0^t (1-\exp(-\lambda w))\ (t-w)^k\,dw$$

By induction one shows the recursive relation

$$\int_0^t \exp(-\lambda w)\ (t-w)^k\, dw$$

$$= t^k/\lambda - k/\lambda \int_0^t \exp(-\lambda w) \times (t-w)^{k-1}\, dw, \qquad k > 0,$$

with starting value

$$\int_0^t \exp(-\lambda w)\, dw = (1-\exp(-\lambda t))/\lambda .$$

It follows that

$$\int_0^t (1-\exp(-\lambda w))\cdot(t-w)^k\, dw = t^{k+1}/(k+1) - k!$$

$$\left[\exp(-\lambda t)/(-\lambda)^{k+1} - \sum_{j=0}^{k} t^j/j!\,(-\lambda)^{k+1-j}\right]$$

Introduced above one obtains

$$I_2 = \exp(-\mu(x+Pt))\cdot[S_1+S_2+S_3]$$

with

$$S_1 = \sum_{k=0}^{\infty} \{1/(k+1)\}\cdot(-\lambda t)^{k+1}/(k+1)!,$$

$$S_2 = -\sum_{k=0}^{\infty} 1/(k+1) \sum_{j=0}^{\infty} (-\lambda t)^j/j!,$$

$$S_3 = \sum_{k=0}^{\infty} 1/(k+1) \sum_{j=0}^{k} (-\lambda t)^j/j!.$$

But one has

$$S_1+S_2+S_3 = -\sum_{k=0}^{\infty} 1/(k+1)\cdot \sum_{j=k+2}^{\infty} (-\lambda t)^j/j!$$

$$= -\sum_{j=2}^{\infty} (-\lambda t)^j/j! \sum_{k=1}^{j-1} 1/k,$$

the last equality being obtained by interchanging the order of summation Therefore formula (4.17) is shown.

ACKNOWLEDGEMENTS

This work was done at the Department of Statistics of the University of Toronto during the academic year 1988/89. The author thanks Prof. D. ANDREWS for the opportunity to teach and make research in Actuarial Science at his Department Furthermore the author is grateful to the referees for pointing out several errors and for making suggestions for improvements.

REFERENCES

BAKER, T H (1977) *The numerical treatment of integral equations* Oxford, Clarendon Press

BELLMAN, R and ROTH, R S (1984) The Laplace Treansform World Scientific

BOWERS, N L, GERBER, H U, HICKMAN, J C, JONES, D A and NESBITT, C J (1986) Actuarial Mathematics Society of Actuaries, Itasca, IL

BÜHLMANN, H (1984) Numerical evaluation of the compound Poisson distribution recursion or fast Fourier transform? *Scand Actuarial J*, 116 126

DE PRIL, N (1985) Recursions for convolutions of arithmetic distributions *ASTIN Bulletin* **15**, 135-139

DE PRIL, N (1986) On the exact computation of the aggregate claims distribution in the individual life model *ASTIN Bulletin* **16**, 109 112

DE PRIL, N and VANDENBROEK, M (1987) Recursions for the distribution of a life portfolio a numerical comparison *Bull of the Royal Association of Belgian Actuaries*

DE PRIL, N (1988) Improved approximations for the aggregate claims distribution of a life insurance portfolio *Scand Actuarial J*

DE PRIL, N (1989) The aggregate claims distribution in the individual model with arbitrary positive claims *ASTIN Bulletin* **19**, 9 24

DOETSCH, G (1976) *Einfuhrung in Theorie und Anwendung der Laplace Transformation*, 3te Auflage Birkhauser Verlag Basel and Stuttgart (English translation (1974) by Springer Verlag)

FELLER, W (1968) *An introduction to probability theory and its applications*, vol 1 and 2 New York John Wiley

GERBER, H U (1979) *An introduction to mathematical Risk Theory* Huebner Foundation for Insurance Education, University of Pennsylvania

GERBER, H U and VALDERRAMA OSPINA, A (1987) A simple proof of Feller's characterization of the compound Poisson distributions *Insurance Mathematics and Economics* **6**, 63-64

GIFFIN, W C (1975) *Transform techniques for probability modeling* Academic Press

HIRSCHMAN, I I and WIDDER, D V (1965) *La transformation de convolution* Paris Gauthier-Villars

HÜRLIMANN, W (1986) Error bounds for stop-loss premiums calculated with the fast Fourier transform *Scand Actuarial J*, 107-113

HÜRLIMANN, W (1989a) On maximum likelihood estimation for count data models *Appears in Insurance Mathematics and Economics*

HÜRLIMANN, W (1989b) *On linear combination of random variables and risk theory* 14 Symposium on Operations Research, Ulm, September 1989

JERRI, A J (1985) *Introduction to integral equations with applications Monographs and textbooks in Pure and Applied Mathematics*, vol 93 New York and Basel Marcel Dekker

KAAS, R, VAN HEERWAARDEN, A E and GOOVAERTS, M J (1988) Between individual and collective model for the total claims *ASTIN Bulletin* **18**

KATTI, S K (1967) Infinite divisibility of integer-valued random variables *Annals of Mathematical Statistics* **38**, 1306 1308

KESTEMONT, R-M and PARIS, J (1985) Sur l'ajustement du nombre de sinistres *Bulletin of the Association of Swiss Actuaries* **85**, 157 164

KESTEMONT, R-M and PARIS, J (1987) *On compound Poisson laws* Paper presented at the meeting on Risk Theory in Oberwolfach, West Germany

PANJER, H H (1981) Recursive evaluation of a family of compound distributions *ASTIN Bulletin* **12**, 22-26

PANJER, H H and WILLMOT, G E (1984) Models for the distribution of aggregate claims in risk theory *Transactions of the Society of Actuaries* **36**, 399–446

REIMERS, L (1988) Letter to the Editor *ASTIN Bulletin* **18**

SCHWARTZ, L (1966) *Théorie des distributions* Nouvelle édition, Hermann, Paris

STEUTEL, F (1970) *Preservation of infinite divisibility under mixing and related results* Math Centre Tracts 33 Amsterdam Mathematical Centre

STEUTEL, F (1979) Infinite divisibility in theory and practice *Scandinavian Journal of Statistics* **6**, 57–64

STRÖTER, B (1985) The numerical evaluation of the aggregate claim density function via integral equations *Blatter der Deutschen Gesellschaft fur Versicherungsmathematik*, 1–13

SUNDT, B and JEWELL, W (1981) Further results on recursive evaluation of compound distributions *ASTIN Bulletin* **12**, 27–39

THYRION, P (1969) Extension of the collective risk theory *Skandinavisk Aktuarietidskrift* **52 (Suppl.)**, 84–98

VAN HARN, K (1978) *Classifying infinitely divisible distributions by functional equations* Math Centre Tracts 103 Amsterdam Mathematical Centre

WIDDER, D V (1971) *An introduction to transform theory* New York and London Academic Press

WILLMOT, G E (1988) Sundt and Jewell's family of discrete distributions *ASTIN Bulletin* **18**, 17–29

WERNER HÜRLIMANN
Allgemeine Mathematik, Winterthur-Leben, CH-8400 Winterthur, Switzerland.

DISTRIBUTIONS IN LIFE INSURANCE

BY JAN DHAENE

Instituut voor Actuariele Wetenschappen, K.U. Leuven, Belgium

ABSTRACT

In most textbooks and papers that deal with the stochastic theory of life contingencies, the stochastic approach is restricted to the computation of expectations and higher order moments. For a wide class of insurances on a single life, we derive the distribution and the probability density function of the benefit and the loss functions. Both the continuous and the discrete case are considered.

KEYWORDS

Single life contingencies, benefit function, loss function, stochastic approach.

1 INTRODUCTION

In the two recent actuarial textbooks of GERBER (1986) and BOWERS et al (1987) the theory of life contingencies is built up in function of the stochastic remaining life time of the insured.

This stochastic approach permits to define two important kinds of stochastic functions for an insurance: the benefit function and the loss function at a certain time. The benefit function of an insured of age x at policy issue is defined as the discounted value of all the benefits to be paid by the insurer over the random future lifetime T_x of the insured. The loss function at time s, given the insured is alive at that time, is the discounted value of all the benefits to be paid by the insurer over the random future lifetime T_{x+s} of the insured less the discounted value of all the premiums to be paid by the insured over the same period

Most results of the traditional deterministic theory are obtained by considering only the expected value of the above defined functions. The net single premium is defined as the expectation of the benefit function The equivalence principle is the requirement that the expected loss at time 0 equals 0. From this requirement the net premiums can be computed The net premium reserve at time s is defined as the expectation of the loss function at time s.

BOWERS et al (1987) state that the probabilistic approach of life contingencies "admits a rich field of random variable concepts such as distribution function, probability density function, expected value, variance and moment generating function" Nevertheless, the literature on this probabilistic approach is mostly restricted to the computation of moments of the benefit and the loss functions, see e.g. POLLARD and POLLARD (1969), WOLTHUIS and VAN HOEK (1984), GERBER (1986) and BOWERS et al (1987)

DE PRIL (1989) gives a survey of the distribution functions (d.f.) and the probability density functions (p.d f.) of the benefit function of most common life insurances and annuities.

In this paper we will consider the benefit and loss functions of a "general insurance", by which we mean a combination of the commonly used life insurances, endowment insurances and life annuities It will be shown that these functions are random variables of a special type The d f. and the p.d.f. of a random variable of this type will then be derived. For completeness both the continuous and discrete case will be treated.

2. CONTINUOUS DESCRIPTION OF SINGLE LIFE CONTINGENCIES

Let $T_x \equiv T$ be a continuous nonnegative random variable representing the future lifetime of a life-aged-x

Using the common actuarial notation, the d.f. of T can be written as

$$(2.1) \qquad F_T(t) = \text{Prob}(T \le t) = \begin{cases} 0 & t < 0 \\ {}_tq_x = 1 - {}_tp_x & t \ge 0 \end{cases}$$

TABLE 1

CONSTANTS FOR THE CONTINUOUS ACTUARIAL FUNCTIONS

L I = Life Insurance, E I = Endowment Insurance, L A = Life Annuity

Name	Notation	a	b	c	m	n
whole L I	$\bar{A}_x$	0	1	0	0	∞
n-year term L I	$\bar{A}^1_{x:\overline{n}\vert}$	0	1	0	0	n
m-year deferred L I	${}_{m\vert}\bar{A}_x$	0	1	0	m	∞
m-year deferred n-year term L I	${}_{m\vert}\bar{A}^1_{x:\overline{n}\vert}$	0	1	0	m	n
n-year pure E I	$\bar{A}_{x:\overset{1}{\overline{n}\vert}}$	0	0	v^n	0	n
n-year E I	$\bar{A}_{x:\overline{n}\vert}$	0	1	v^n	0	n
m-year deferred n-year E I	${}_{m\vert}\bar{A}_{x:\overline{n}\vert}$	0	1	v^{m+n}	m	n
whole L A	$\bar{a}_x$	$1/\delta$	$-1/\delta$	0	0	∞
n-year temporary L A	$\bar{a}_{x:\overline{n}\vert}$	$1/\delta$	$-1/\delta$	$\bar{a}_{\overline{n}\vert}$	0	n
m-year deferred whole L A	${}_{m\vert}\bar{a}_x$	v^m/δ	$-1/\delta$	0	m	∞
m-year deferred n-year temporary L A	${}_{m\vert}\bar{a}_{x:\overline{n}\vert}$	v^m/δ	$-1/\delta$	$v^m\,\bar{a}_{\overline{n}\vert}$	m	n

with ${}_0q_x = 0$ and $\lim_{t \to \infty} {}_tq_x = 1$.

The p d.f of T is given by

$$(2\ 2) \qquad f_T(t) = F'_T(t) = \begin{cases} 0 & t < 0 \\ {}_tp_x \ \ \mu_{x+t} & t \geq 0 \end{cases}$$

where μ_x denotes the force of mortality of a life aged x

From Table 1 it can be seen that the benefit function of the common life insurances, endowment insurances and life annuities on a single life aged x at policy issue can be written as a stochastic variable of the form

$$(2.3) \qquad S = \begin{cases} 0 & 0 \leq T < m \\ a + bv^T & m \leq T < m+n \\ c & T \geq m+n \end{cases}$$

where a, b and c are real numbers and m and n are nonnegative integers. Further, $v = 1/(1+i)$ is the present value factor related to the annual valuation rate of interest i.

In Table 1 the following notation is used. $\delta = \ln(1+i)$ is the force of interest associated with the valuation rate of interest i and $\bar{a}_{\overline{n}|} = (1-v^n)/\delta$ is a continuous n-year temporary annuity certain

A general continuous insurance on a single life aged x at policy issue is defined as a combination of the life insurances, endowment insurances and life annuities considered in Table 1 and where the premiums are paid by a combination of the life annuities and pure endowment insurances of Table 1

The stochastic variable describing the benefit function of a general continuous insurance is a linear combination of random variables of the form (2 3) So it follows immediately that this stochastic variable can be written as

$$(2.4) \qquad S = a_i + b_i v^T \qquad m(i-1) \leq T < m(i), \qquad i = 1, \ \ , n$$

with $T \equiv T_x$, a_i and b_i $(i = 1, \ldots, n)$ real numbers and $m(i)$ $(i = 0, \ldots, n)$ nonnegative integers satisfying

$$(2.5) \qquad 0 \leq m(0) < m(1) < \quad < m(n) \leq \infty$$

It is easy to see that the loss function at times s $(s \geq 0)$, given survival of the insured at that time, can also be described by a stochastic variable of the form (2.4) with $T \equiv T_{x+s}$

The p.d.f of a random variable of the form (2.4) will be derived in the following theorem The delta-function will be denoted by $\Delta(x)$ to avoid confusion with the symbol δ for the force of interest. For a study of the delta-function see e g. PAPOULIS (1962)

Theorem 1. Let S be the stochastic variable defined in (2 4) with $T \equiv T_z$ Define for $i = 1, \ldots, n$

$$(2.6) \qquad m(i)^- = a_i + \min\left(b_i v^{m(i-1)}, b_i v^{m(i)}\right)$$

$$(2.7) \qquad m(i)^+ = a_i + \max\left(b_i v^{m(i-1)}, b_i v^{m(i)}\right)$$

The p.d.f. of S is given by

$$(2.8) \qquad f(s) = \sum_{i=1}^{n} G_i(s)$$

with for $i = 1, \ldots, n$

$$(2.9) \qquad G_i(s) = \begin{cases} \Delta(s-a_i)\left({}_{m(i-1)}p_z - {}_{m(i)}p_z\right) & : b_i = 0 \\ {}_{r(i)}p_z \; \mu_{z+r(i)}/|\delta(s-a_i)| & : b_i \neq 0 \text{ and } m(i)^- < s < m(i)^+ \\ 0 & \text{elsewhere} \end{cases}$$

where $r(i)$ is given by

$$(2.10) \qquad r(i) = -\frac{1}{\delta} \ln\left(\frac{s-a_i}{b_i}\right) \qquad \frac{s-a_i}{b_i} > 0$$

Proof. Using the Law of Total Probability the p d f. of S can be written in the form (2 8) with

$$G_i(s) = f(s \mid m(i-1) \leq T < m(i)) \operatorname{Prob}(m(i-1) \leq T < m(i))$$

For $b_i = 0$ it follows that

$$f(s \mid m(i-1) \leq T < m(i)) = \Delta(s-a_i).$$

Consider now the case $b_i \neq 0$ We obtain

$$G_i(s) = \begin{cases} f_T(r(i))/(\delta(s-a_i)) & b_i > 0,\ a_i + b_i v^{m(i)} < s < a_i + b_i v^{m(i-1)} \\ f_T(r(i))/(\delta(a_i - s)) & b_i < 0,\ a_i + b_i v^{m(i-1)} < s < a_i + b_i v^{m(i)} \\ 0 & : \text{elsewhere}. \end{cases}$$

with $r(i)$ defined in (2 10)

So it follows that $G_i(s)$ is given by (2 9).

The d f of S is derived in the next theorem The following notation will be used:

$$(x)_+ = \max(0, x)$$

and

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Theorem 2. The d.f of the random variable S defined in (2 4) with $T \equiv T_z$ is given by

$$F(s) = \sum_{i=1}^{n} \left({}_{\alpha(i)}q_z - {}_{\beta(i)}q_z\right)_+ K_i(s) \tag{2 12}$$

where $\alpha(i)$ and $\beta(i)$ $(i = 1, \dots, n)$ are given by

$$\alpha(i) = \begin{cases} \max\left\{0, \min\left\{m(i), -\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)\right\}\right\} & . \ b_i < 0, \ s < a_i \\ m(i) & . \text{ elsewhere} \end{cases} \tag{2 13}$$

$$\beta(i) = \begin{cases} m(i-1) & . \ b_i \le 0 \\ \max\left\{m(i-1), -\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)\right\} & b_i > 0, \ s > a_i \\ m(i) & b_i > 0, \ s \le a_i \end{cases} \tag{2 14}$$

Finally, $K_i(s)$ $(i = 1, \dots, n)$ is defined as

$$K_i(s) = \begin{cases} H(s-a_i) & : \ b_i = 0 \\ 1 & b_i \ne 0 \end{cases} \tag{2 15}$$

Proof. Using the Law of Total Probability we find that

$$F(s) = \sum_{i=1}^{n} \text{Prob}\left(a_i + b_i v^T \le s \text{ and } m(i-1) \le T < m(i)\right) \tag{2.16}$$

It follows that

$$\text{Prob}\left(a_i + b_i v^T \le s \text{ and } m(i-1) \le T < m(i)\right)$$
$$= \begin{cases} H(s-a_i)\,\text{Prob}\left(m(i-1) \le T < m(i)\right) & \cdot \ b_i = 0 \\ \text{Prob}\left(\max\left\{m(i-1), -\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)\right\} \le T < m(i)\right) & . \ b_i > 0, \ s > a_i \\ 0 & \cdot \ b_i > 0, \ s \le a_i \\ \text{Prob}\left(m(i-1) \le T < m(i)\right) & . \ b_i < 0, \ s \ge a_i \\ \text{Prob}\left(m(i-1) \le T < \min\left\{m(i), -\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)\right\}\right) & \cdot \ b_i < 0, \ s < a_i \end{cases}$$

Or

$$\text{Prob}\left(a_i + b_i v^T \le s \text{ and } m(i-1) \le T < m(i)\right) = \left({}_{\alpha(i)}q_z - {}_{\beta(i)}q_z\right)_+ K_i(s)$$

with $\alpha(i)$ and $\beta(i)$ defined in (2.13) and (2.14).
Now (2.12) is obtained with the help of (2.16)

The p.d.f. and the d.f. of the benefit function of the insurances and annuities considered in Table 1 can be written in a simpler form as is proven in the following corollary.

Corollary 1. Let S be the stochastic variable defined in (2.3) with $T \equiv T_x$.
Define

$$m^- = a + \min(b\,v^m,\ b\,v^{m+n}) \tag{2.17}$$

$$m^+ = a + \max(b\,v^m,\ b\,v^{m+n}) \tag{2.18}$$

The p.d.f. of S is given by

$$f(s) = {}_mq_x\,\Delta(s) + G(s) + {}_{m+n}p_x\,\Delta(s-c) \tag{2.19}$$

where $G(s)$ is defined as

$$G(s) = \begin{cases} \Delta(s-a)\ ({}_mp_x - {}_{m+n}p_x) & : b = 0 \\ {}_rp_x\ \mu_{x+r}/|\delta(s-a)| & : b \neq 0 \text{ and } m^- < s < m^+ \\ 0 & : \text{elsewhere} \end{cases} \tag{2.20}$$

with

$$r = -\frac{1}{\delta}\ln\left(\frac{s-a}{b}\right) \qquad \frac{s-a}{b} > 0 \tag{2.21}$$

The d.f. of S is given by

$$F(s) = {}_mq_x\,H(s) + ({}_\alpha q_x - {}_\beta q_x)_+\,K(s) + {}_{m+n}p_x\,H(s-c) \tag{2.22}$$

with

$$\alpha = \begin{cases} \max\left\{0,\ \min\left\{m+n,\ -\frac{1}{\delta}\ln\left(\frac{s-a}{b}\right)\right\}\right\} & : b < 0,\ s < a \\ m+n & \text{elsewhere} \end{cases} \tag{2.23}$$

$$\beta = \begin{cases} m & : b \leq 0 \\ \max\left\{m,\ -\frac{1}{\delta}\ln\left(\frac{s-a}{b}\right)\right\} & b > 0,\ s > a \\ m+n & : b > 0,\ s \leq a \end{cases} \tag{2.24}$$

$$(2.25)\quad K(s) = \begin{cases} H(s-a) & : b = 0 \\ 1 & : b \neq 0 \end{cases}$$

Proof. The random variable S defined in formula (2.3) is a special case of the random variable defined by (2.4) with the constants $n = 3$, $a_1 = 0$, $b_1 = 0$, $a_2 = a$, $b_2 = b$, $a_3 = c$, $b_3 = 0$, $m(0) = 0$, $m(1) = m$, $m(2) = m+n$, $m(3) = \infty$ and $T \equiv T_x$.

Using Theorems 1 and 2, after some straightforward calculation one obtains formulae (2.19) and (2.22).

The d.f. and the p.d.f. of all the continuous insurances and annuities considered in DE PRIL (1989) can be obtained by using Table 1 and Corollary 1

TABLE 2

CONSTANTS FOR THE DISCRETE ACTUARIAL FUNCTIONS

L I = Life Insurance, E I = Endowment Insurance, L A = Life Annuity

Name	Notation	a	b	c	m	n
whole L I	A_x	0	1	0	0	∞
n-year term L I	$A^1_{x:\overline{n}\vert}$	0	1	0	0	n
m-year deferred L I	$_{m}A_x$	0	1	0	m	∞
m-year deferred n-year term L I	$_{m\vert}\bar{A}^1_{x:\overline{n}\vert}$	0	1	0	m	n
n-year pure E I	$A_{x:\overline{n}\vert}^{\;\;1}$	0	0	v^n	0	n
n-year E I	$A_{x:\overline{n}\vert}$	0	1	v^n	0	n
m-year deferred n-year E I	$_{m\vert}A_{x:\overline{n}\vert}$	0	1	v^{m+n}	m	n
whole L A due	a_x	$1/d$	$-1/d$	0	0	∞
whole L A immediate	a_x	$1/i$	$-1/d$	0	0	∞
n-year temporary L A due	$a_{x:\overline{n}\vert}$	$1/d$	$-1/d$	$a_{\overline{n}\vert}$	0	n
n-year temporary L A immediate	$a_{x:\overline{n}\vert}$	$1/i$	$-1/d$	$a_{\overline{n}\vert}$	0	n
m-year deferred whole L A due	$_{m}a_x$	v^m/d	$-1/d$	0	m	∞
m-year deferred whole L A immediate	$_{m\vert}a_x$	v^m/i	$-1/d$	0	m	∞
m-year deferred n-year temporary L A due	$_{m}a_{x:\overline{n}\vert}$	v^m/d	$-1/d$	$v^m a_{\overline{n}\vert}$	m	n
m-year deferred n-year temporary L A immediate	$_{m\vert}a_{x:\overline{n}\vert}$	v^m/i	$-1/d$	$v^m a_{\overline{n}\vert}$	m	n

3. DISCRETE DESCRIPTION OF SINGLE LIFE CONTINGENCIES

Let $K \equiv K_x$ be a nonnegative random variable, representing the number of full years to death of a life-aged-x

The distribution of K can then be written as

$$(3.1) \qquad F_K(k) = \text{Prob}\,(K \leq k) = {}_{k+1}q_x = 1 - {}_{k+1}p_x \qquad k = 0, 1, 2, \ldots$$

with $\lim_{k \to \infty} {}_kq_x = 1$

The p.d.f of K is given by

$$(3.2) \qquad f_K(k) = {}_kp_x - {}_{k+1}p_x = {}_{k|}q_x \qquad k = 0, 1, 2, \ldots$$

The benefit functions of the common discrete life insurances, endowment insurances and life annuities on a single life aged x at policy issue can be defined as stochastic variables of the form

$$(3.3) \qquad S = \begin{cases} 0 & : K = 0, 1, \ldots, m-1 \\ a + b\, v^{K+1} & : K = m, m+1, \ldots, m+n-1 \\ c & : K = m+n, m+n+1, \ldots \end{cases}$$

The suiting values for a, b, c, m and n are given in Table 2

The following notation is used. $d = 1 - v$, $\ddot{a}_{\overline{n}|} = (1 - v^n)/d$ is a discrete n-year temporary annuity due and $a_{\overline{n}|} = (1 - v^n)/i$ is a discrete n-year temporary annuity immediate.

A general discrete insurance on a single life-aged-x is defined as a combination of the insurances defined in Table 2 The premiums are paid by a combination of the life annuities and pure endowment insurances of Table 2.

The benefit function and the loss function of a general discrete insurance can be described by a stochastic variable S of the form

$$(3.4) \qquad S = a_i + b_i\, v^{K+1} \qquad m(i-1) \leq K < m(i); \qquad i = 1, \ldots, n$$

with $K \equiv K_x$ for the benefit function and $K \equiv K_{x+s}$ for the loss function at time s. Further, a_i and b_i $(i = 1, \ldots, n)$ are real numbers and $m(i)$ $(i = 0, 1, \ldots, n)$ are nonnegative integers satisfying

$$(3.5) \qquad 0 \leq m(0) < m(1) < \ldots < m(n) \leq \infty$$

In the following theorems the p.d.f. and the d.f. of S are derived.

Theorem 3. The p.d.f. of the variable S defined in (3.4) where $K \equiv K_z$ is given by

$$(3.6) \qquad f(s) = \sum_{i=1}^{n} G_i(s)$$

where for $i = 1, \ldots, n$ the functions $G_i(s)$ are given by

$$(3.7) \qquad G_i(s) = \begin{cases} ({}_{m(i-1)}p_z - {}_{m(i)}p_z)\, \Delta(s-a_i) & : b_i = 0 \\ \sum_{k=m(i-1)}^{m(i)-1} {}_{k|}q_z\, \Delta(s-a_i-b_i v^{k+1}) & : b_i \neq 0 \end{cases}$$

Proof. For $m(i-1) \leq K < m(i)$ and $b_i = 0$ we get that $S = a_i$ or

$$f(s \mid m(i-1) \leq K < m(i)) = \Delta(s-a_i)$$

If $m(i-1) \leq K < m(i)$ and $b_i \neq 0$ the possible values for S are

$$a_i + b_i v^{k+1} \qquad k = m(i-1), \ldots, m(i)-1$$

with respective probabilities

$$ {}_{k|}q_x / \text{Prob}\,(m(i-1) \leq K < m(i))$$

So we find for $b_i \neq 0$

$$f(s \mid m(i-1) \leq K < m(i))\ \text{Prob}\,(m(i-1) \leq K < m(i))$$
$$= \sum_{k=m(i-1)}^{m(i)-1} {}_{k|}q_x\, \Delta(s-a_i-b_i v^{k+1})$$

By using the Law of Total Probability we obtain formula (3.6).

Theorem 4. The d f. of the random variable S defined in (3 4) where $K \equiv K_z$ is given by

$$(3\ 8) \qquad F(s) = \sum_{i=1}^{n} ({}_{\alpha(i)}q_z - {}_{\beta(i)}q_z)_+\, K_i(s)$$

where for $i = 1, \ldots, n$ the functioins $\alpha(i)$, $\beta(i)$ and $K_i(s)$ are given by

$$(3.9) \qquad \alpha(i) = \begin{cases} \max\left\{0, \min\left\{m(i), \left[-\frac{1}{\delta} \ln\left(\frac{s-a_i}{b_i}\right)\right]\right\}\right\} & : b_i < 0,\ s < a_i \\ m(i) & \text{elsewhere} \end{cases}$$

$$(3\ 10) \qquad \beta(i) = \begin{cases} m(i-1) & b_i \leq 0 \\ \max\left\{m(i-1), \left]-\frac{1}{\delta} \ln\left(\frac{s-a_i}{b_i}\right) - 1\right[\right\} & : b_i \leq 0,\ s > a_i \\ m(i) & : b_i > 0,\ s \leq a_i \end{cases}$$

$$(3.11)\qquad K_i(s) = \begin{cases} H(s-a_i) & : b_i = 0 \\ 1 & : b_i \neq 0 \end{cases}$$

For a real number x, $]x[$ denotes the smallest integer greater than or equal to x and $[x]$ denotes the greatest integer less than or equal to x.

Proof. For $m(i-1) \leq K < m(i)$ we find

Prob $(a_i + b_i\, v^{K+1} \leq s$ and $m(i-1) \leq K < m(i))$

$$= \begin{cases} H(s-a_i)\,\text{Prob}\,(m(i-1) \leq K < m(i)) & : b_i = 0 \\ \text{Prob}\left(\max\left\{m(i-1), \left]-\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)-1\right[\right\} \leq K < m(i)\right) & : b_i > 0, s > a_i \\ 0 & : b_i > 0, s \leq a_i \\ \text{Prob}\,(m(i-1) \leq K < m(i)) & : b_i < 0, s \geq a_i \\ \text{Prob}\left(m(i-1) \leq K < \min\left\{m(i), \left[-\frac{1}{\delta}\ln\left(\frac{s-a_i}{b_i}\right)\right]\right\}\right) & : b_i < 0, s < a_i \end{cases}$$

Or

$$\text{Prob}\,(a_i + b_i\, v^{K+1} \leq s \text{ and } m(i-1) \leq K < m(i)) = ({}_{\alpha(i)}q_z - {}_{\beta(i)}q_z)_+\; K_i(s)$$

with $\alpha(i)$, $\beta(i)$ and $K_i(s)$ defined in (3.9), (3.10) and (3.11).
By using the Law of Total Probability we obtain the desired result.

The p.d.f and the d.f. of the benefit function of the discrete insurances and annuities considered in Table 2 can be written in a simpler form which is derived in the next corollary.

Corollary 2. Let S be the stochastic variable defined in (3.3) with $K \equiv K_x$.
The p.d.f. of S is given by

$$(3.12)\qquad f(s) = {}_mq_x\,\Delta(s) + G(s) + {}_{m+n}p_x\,\Delta(s-c)$$

with

$$(3.13)\qquad G(s) = \begin{cases} ({}_mp_x - {}_{m+n}p_x)\,\Delta(s-a) & : b = 0 \\ \sum_{k=m}^{m+n-1} {}_{k|}q_x\,\Delta(s-a-b\,v^{k+1}) & : b \neq 0 \end{cases}$$

The d.f of S is given by

$$(3.14)\qquad F(s) = {}_mq_x\,H(s) + ({}_\alpha q_x - {}_\beta q_x)_+\,K(s) + {}_{m+n}p_x\,H(s-c)$$

with

$$(3.15)\quad \alpha = \begin{cases} \max\left\{0,\ \min\left\{m+n, \left[-\frac{1}{\delta}\ln\left(\frac{s-a}{b}\right)\right]\right\}\right\} & : b < 0,\ s < a \\ m+n & : \text{elsewhere} \end{cases}$$

$$(3.16)\quad \beta = \begin{cases} m & : b \le 0 \\ \max\left\{m, \left]-1-\frac{1}{\delta}\ln\left(\frac{s-a}{b}\right)\right[\right\} & : b > 0,\ s > a \\ m+n & : b > 0,\ s \le a \end{cases}$$

$$(3.17)\quad K(s) = \begin{cases} H(s-a) & : b = 0 \\ 1 & : b \neq 0 \end{cases}$$

Proof. The proof follows immediately from Theorems 3 and 4.

The p d f and the d f. of the discrete insurances and annuities considered in De Pril (1989) can be derived with the help of Table 2 and Corollary 2.

4. EXAMPLE

A person aged x purchases a combination benefit consisting of a n-year term life insurance of I payable immediately on his death and a n-year deferred whole life annuity of J per annum payable continuously while he survives beyond age $x+n$

Let the benefit functions of the insurances and annuities defined in Table 1 be denoted by adding a tilde to the usual deterministic symbols. The benefit function of the continuous general insurance defined above is then given by

$$S = I\,\tilde{\bar{A}}^{1}_{x\,\overline{n}|} + J\,{}_{n}\tilde{\bar{a}}_{x}$$

By using (2.3) and Table 1 this benefit function can be written as a variable of the form (2.4) with $T \equiv T_x$:

$$S = \begin{cases} I\,v^T & : 0 \le T < n \\ J\,\frac{v^n - v^T}{\delta} & : T \ge n \end{cases}$$

From Theorem 1 it follows that the p.d f of S is given by

$$f(s) = G_1(s) + G_2(s)$$

with

$$G_1(s) = \begin{cases} {}_{r(1)}p_x \ \mu_{x+r(1)}/(\delta s) & : Iv^n < s < I \\ 0 & : \text{elsewhere} \end{cases}$$

$$G_2(s) = \begin{cases} {}_{r(2)}p_x \ \mu_{x+r(2)}/(Jv^n - \delta s) & : 0 < s < Jv^n/\delta \\ 0 & : \text{elsewhere} \end{cases}$$

and

$$r(1) = -(1/\delta) \ln (s/I)$$

$$r(2) = -(1/\delta) \ln (v^n - (\delta s)/J)$$

The d.f. of S follows from Theorem 2

$$F(s) = ({}_nq_x - {}_\beta q_x)_+ + ({}_aq_x - {}_nq_x)_+$$

with

$$a = \begin{cases} \max\{0, \ r(2)\} & s < J\,v^n/\delta \\ \infty & s \geq J\,v^n/\delta \end{cases}$$

$$\beta = \begin{cases} n & : s \leq 0 \\ \max\{0, \ r(1)\} & : s > 0 \end{cases}$$

REFERENCES

BOWERS, N L, GERBER, H U, HICKMAN, J C, JONES, D A and NESBITT, C J (1987) *Actuarial Mathematics* Society of Actuaries, Itacsa, IL

DE PRIL, N (1989) The distributions of Actuarial functions *Bulletin of the Swiss Association of Actuaries*, 173 183

GERBER, H U (1986) *Lebensversicherungsmathematik* Vereinigung schweizerischer Versicherungs-mathematiker, Zurich

PAPOULIS, A (1962) *The Fourier Integral and its applications* Mc Graw-Hill, New York

POLLARD, A H and POLLARD, J H (1969) A stochastic approach to actuarial functions *Journal of the Institute of Actuaries*, 79-113

WOLTHUIS, H and VAN HOEK, I (1984) Stochastic models for life contingencies *Insurance Mathematics and Economics* 5, 217-254

JAN DHAENE

Instituut voor Actuariele Wetenschappen, K U. Leuven, Dekenstraat 2, B-3000 Leuven, Belgium.

PREDICTING IBNYR EVENTS AND DELAYS

II. Discrete Time

By William S. Jewell

Engineering Systems Research Center
University of California at Berkeley

Abstract

An IBNYR event is one that occurs randomly during some fixed exposure interval and incurs a random delay before it is reported. A previous paper developed a continuous-time model of the IBNYR process in which both the Poisson rate at which events occur and the parameters of the delay distribution are unknown random quantities; a full-distributional Bayesian method was then developed to predict the number of unreported events Using a numerical example, the success of this approach was shown to depend upon whether or not the occurrence dates were available in addition to the reporting dates. This paper considers the more usual practical situation in which only discretized epoch information is available, this leads to a loss of predictive accuracy, which is investigated by considering various levels of quantization for the same numerical example.

Keywords

Incurred But Not Reported (IBNR) models; reporting delays; Bayesian estimation and prediction; Bayesian approximations; discrete-time models.

1. Introduction

An Incurred But Not Yet Reported claim in insurance is an event whose occurrence during some fixed exposure interval is not known until some later date because of random reporting delays. These claims, plus the Incurred But Not Fully Reported claims, which have been reported but whose cost development is incomplete, form the Incurred But Not Reported (IBNR) portfolio for a given policy exposure interval The accurate prediction of the total number and the ultimate costs of such claims is a critical and recurring problem in many insurance lines

In Jewell (1989), hereinafter referred to as IBNYR-I, the author developed a continuous-time model for predicting the number of unreported IBNYR events, under the assumptions that the random (Poisson) rate of event occurrence as well as the parameters of the delay distribution are unknown

Examination of the likelihood revealed not only a coupling between the unknown parameters for the number of occurrences and their associated random delays, but a strong dependence upon the type of epoch data available, for example, having only reporting dates but not occurrence dates led to predictions with wider variances than when both dates were available A Bayesian development was then used to obtain a full predictive distribution and, from it, the interesting point predictors; natural conjugate priors were used for simplicity, although extensions to empirical priors are immediate. Either way, the key computational issue is the evaluation of the ratio of two integrals, for which various good approximation techniques are available. So predictive means, variances, and tail probabilities for IBNYR events are now easily obtained under continuous-time assumptions

However, in most firms, exact epoch data is difficult to obtain, is unreliable, or, possibly, is dismissed as being unimportant For instance, most models in the IBNR literature use quantized reporting intervals that are one year long, the same length as the usual exposure period While this may give satisfactory results for the long-duration *cost evolution* of many casualty claims, *reporting delays* may be shorter than or comparable to the exposure interval, so that gross discretization can, as we shall see, lead to a significant loss in predictive power. Exceptions might be claims for industrial diseases (such as asbestosis) or for product liability, both of which may take a long time to develop

The model we develop below is parallel to that of IBNYR-I, except that the reporting of dates is discretized into intervals equal to, or a submultiple of, the basic exposure interval We model the equivalents of the first two cases of epoch data described in IBNYR-I (reporting dates always observed, occurrence dates may or may not be reported), since we know that both classical and Bayesian predictions are already bad in the other continuous cases where only occurrence dates, or only counts-to-date are available To compare the effects of changing from continuous to quantized data, we consider the same numerical example as in the first paper

Important references on the IBNR problem were given in IBNYR-I; supplemented by those below, they together give an overview of research in this area, most of which emphasizes point estimates for discrete-time cost-evolution models. Our results will not parallel these other efforts until a planned third paper on the "IBNR triangle" appears, in which the effect of *collateral discretized* data from several exposure periods is analyzed. As discussed in IBNYR-I, we believe that it is important to understand thoroughly the effect of various modelling assumptions upon event prediction before adding on the dynamics of random cost evolution

2 THE MODEL

As in IBNYR-I, we assume that, during an *exposure interval* $(0, T]$, a random number of events, $\tilde{n}$, occurs according to a Poisson process with parameter λT This implies that, given $\tilde{n} = n$, the *occurrence epochs* $(\tilde{x}_1, \tilde{x}_2, \ldots \tilde{x}_n)$ of the events are, *a priori*, independent and uniformly distributed over $(0, T]$.

Associated with each event indexed k is a random *reporting delay*, $\tilde{w}_k$, so that the actual *observation* or *reporting epochs* are $\tilde{y}_k = \tilde{x}_k + \tilde{w}_k$ $(k = 1, 2, \ldots n)$ Each delay is assumed to be i.i.d with a common probability density, $f(w \mid \theta)$, that depends upon one or more parameters, θ. It follows that, is given θ, each event pair $(\tilde{x}_k, \tilde{y}_k)$ is i.i.d. with joint density:

$$(2.1) \qquad p(x, y \mid \theta) = \frac{1}{T} f(y - x \mid \theta) \qquad (0 \leq x \leq T,\ x \leq y \leq \infty)$$

over the semi-infinite wedge-shaped region shown in Figure 1, and zero elsewhere. If we observe the reporting dates of the IBNYR events over some *observation interval* $(0, t]$, it is clear that only those pairs with $y_k \leq t$ will actually be reported, so that the total number of *reported events* will be some number R less than n

As before, we assume that λ and θ are outcomes of the unknown random quantities $\tilde{\lambda}$ and $\tilde{\theta}$, respectively, for convenience *a priori* independent with known *prior densities*, $p(\lambda)$ and $p(\theta)$. Suppose that *epoch data* $\mathcal{D}_k$ is observed for each of the R reported events. Given these priors and the *total data*, $\mathcal{D} = \{R, \bigcup \mathcal{D}_k\}$, the *parameter estimation problem* is to determine $p(\lambda, \theta \mid \mathcal{D})$ and the *event prediction problem* is to determine $p(u \mid \mathcal{D})$, where $\tilde{u} = \tilde{n} - R$ is the unknown number of *unobserved IBNYR events* still outstanding

To introduce the effects of discrete-time reporting, we imagine that the time axis is partitioned into equal *reporting intervals*, $\mathcal{I}_l = ((l-1)\Delta, l\Delta]$ $(l = 1, 2, \ldots)$, thus $\Delta \leq T$ is the common length of the reporting intervals, and the precise values of any dates within that interval are lost. We assume that Δ is a submultiple of T, so that $I = T/\Delta$, the *quantization level*, is a positive integer In practice, T is usually one year, and $I = 1, 2, 4$, or 12. The observation interval $(0, t]$ can now only be, say, $t = J\Delta$, with $J = 1, 2, \ldots$

We now consider two cases of quantized epoch reporting that correspond to the continuous data types I and II analyzed in IBNYR-I.

2.1. Type Iq Data. Quantized Occurrence and Reporting Dates

In this case, the continuous-time epoch data (x_k, y_k) for an observed event indexed k is mapped into $\mathcal{D}_k = (i_k, j_k)$, two positive integers indicating the reporting intervals, viz $(i, j) \equiv (x \in \mathcal{I}_i) \bigcap (y \in \mathcal{I}_j)$. Obviously, $(1 \leq i \leq I)$ and $(j \geq i)$ always Figure 1, which shows the joint partitioning of the allowed region for $I = 4$ and $t = 4.0$, gives a "tiling" that helps us to visualize the quantization Most of the tiles are squares with sides Δ, but, if x and y are in the same interval, then (j, j) is reported in a triangular region, since $x \leq y$ always

The probabilities associated with each tile can be expressed most easily with the aid of the function:

$$(2.2) \qquad \Phi_h(\theta) = \frac{1}{T} \int_{(h-1)\Delta}^{h\Delta} F(w \mid \theta)\, dw \qquad (h = 1, 2, \ldots)$$

$(\Phi_0(\theta) = 0)$, which is monotonic over the integers and approaches I^{-1} for

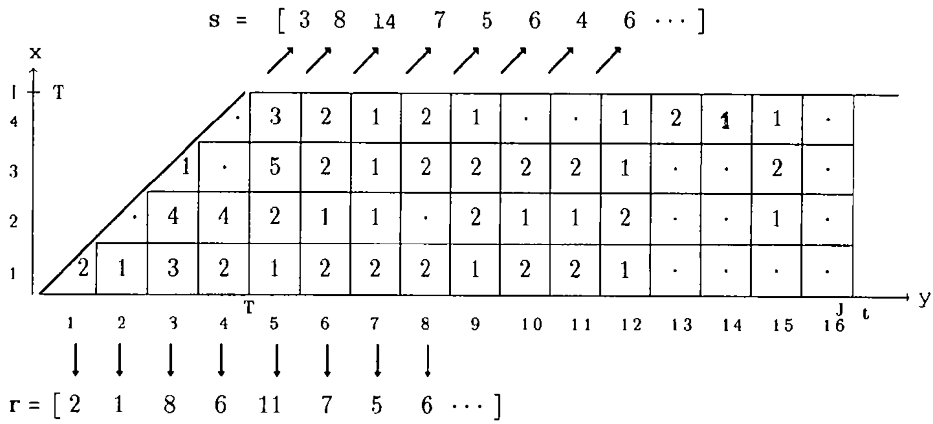


FIGURE 1 Regions of definition of quantized occurence and reporting dates, showing the distribution of 74 of the 100 events generated with $\theta = 0.5$ year^{-1}, for $I = 4$ and $J = 16$

every θ as $h \to \infty$. Letting $\pi_{ij}(\theta) = \mathcal{P}\{\ell_k = (i,j)|\theta\}$ (any k) be the mass associated with tile (i,j), we find from (2 1) that:

$$\pi_{ij}(\theta) = \Phi_{j-i+1}(\theta) - \Phi_{j-i}(\theta) \qquad (1 \le i \le I)(j \ge i). \tag{2 3}$$

In other words, the mass of each cell along the "diagonals" with constant $h = j-i+1$ ($h = 1, 2, \ldots$) is the same, which might be expected from first principles This is the discrete equivalent of a likelihood that depends only on $w = y-x$ ($w > 0$), as in the Type I continuous-time data models, in fact, if $(j-i) = w$ and $\Delta \to 0$, (2 3) approaches $f(w|\theta)\Delta^2/T$, so that events with about the same w carry the same information in the limit

Suppose a total of R events were reported during the observation interval $(0, J\Delta]$; this includes only events for which $j \le t/\Delta$ Rather than reporting the discrete dates (i,j) for each event k, we can imagine that the epoch data represents a *distribution* of the R events into r_{ij} events for each tile (i,j), following a multinomial law with probabilities equal to $\pi_{ij}(\theta)$, normalized by dividing by the sum of probabilities over all cells in the observation interval. However, because of the structure of (2.3), the $\{r_{ij}\}$ can be accumulated over cells of equal mass on each diagonal, reducing them to the *sufficient statistics for Type Iq data*:

$$s_h = \sum_{i=1}^{\min(I,\, J+1-h)} r_{i,\, i+h-1} \qquad (h = 1, 2, \ldots, J) \tag{2 4}$$

The complicated upper limit restricts the length of the observable "diagonal" elements as h approaches J and if $J < I$.

Figure 1 shows how the 74 counts for $J = 16$ and $I = 4$ in the numerical example are distributed over the cells. We find easily that $s = [3, 8, 14, 7, 5, 6, 4, 6, 5, 6, 5, 2, 2, \underline{1}, \underline{0}, \underline{0}]^T$, but note that, because of (2.4), if we decide to increase J, then the last (underlined) $I-1$ numbers would have to be increased by any new counts on their diagonals!

If we express the probabilities (2.3) in terms of

$$\varphi_h(\theta) = \Phi_h(\theta) - \Phi_{h-1}(\theta) \qquad (h = 1, 2, \ldots), \tag{2.5}$$

the multinomial *conditional data likelihood*, given R and θ, is

$$p(\cup \mathscr{Y}_k \mid R, \theta) = \binom{R}{\mathbf{s}} \prod_{h=1}^{J} [\varphi_h(\theta)]^{s_h} \Big/ \left[\sum_{l=1}^{I} \min(I, J+1-l)\, \varphi_l(\theta) \right]^R, \tag{2.6}$$

where $\mathbf{s} = [s_1, s_2, \ldots s_J]^T$ is defined over the discrete simplex, $0 \le s_h \le R$, $\Sigma s_h = R$. Note how the total normalizing mass requires a weighted sum of all the $\{\varphi_h(\theta)\}$ to account for the fewer tiles near $h = J$.

2.2. Type IIq Data. Quantized Reporting Dates

The situation is somewhat simpler with only reporting epochs, $\mathscr{Y}_k = (j_k)$, given for each event, which means that all event counts and probabilities are merged in each "column" of cells in Figure 1. Thus, the *sufficient statistics for Type IIq data* are $\mathbf{r} = [r_1, r_2, \ldots r_J]^T$, where:

$$r_j = \sum_{i=1}^{\min(I, j)} r_{ij} \qquad (j = 1, 2, \ldots J). \tag{2.7}$$

This gives $\mathbf{r} = [2, 1, 8, 6, 11, 7, 5, 6, 6, 5, 5, 5, 2, 1, 4, 0]^T$ from Figure 1. (2.7) can also be thought of as the result of a multinomial sorting of R events, this time with probabilities

$$\pi_j(\theta) = \sum_{i=1}^{\min(j, I)} \pi_{ij}(\theta) = \Phi_j(\theta) - \Phi_{j-I}(\theta) \qquad (j = 1, 2, \ldots), \tag{2.8}$$

where the second term vanishes if $j \le I$.

Thus, for Type IIq epoch data, (2.6) is replaced by:

$$p(\cup \mathscr{Y}_k \mid R, \theta) = \binom{R}{\mathbf{r}} \prod_{j=1}^{J} [\pi_j(\theta)]^{r_j} \Big/ \left[\sum_{l=1}^{J} \pi_l(\theta) \right]^R, \tag{2.9}$$

with $\mathbf{r}$ defined over the discrete R-simplex. Here the normalizing mass is simpler because each $\pi_j(\theta)$ is already the sum of individual tile probabilities in each column.

As $\Delta \to 0$, (2.8) reduces to Δ times the usual probability for continuous Type II data, that is, $[F(t \mid \theta) - F((t-T)^+ \mid \theta)]\, \Delta / T$. Of course, when $I = 1$ and $\Delta = T$, the distinction between discrete Cases Iq and IIq vanishes, since $s_j = r_j = r_{1j}$, and $\varphi_j(\theta) = \pi_j(\theta) = \pi_{1j}(\theta)$.

3. DATA LIKELIHOODS AND MLE ESTIMATES

In the next two sections, we assume that Type IIq data is available, however, all formulae in which $\{r_j\}$ and $\{\pi_j(\theta)\}$ are used can be changed to Type Iq simply by replacing them with $\{s_h\}$ and $\{\varphi_h(\theta)\}$, respectively Our first step is to uncondition (2.6) and (2.9) on R by noting that, given n and θ, n can be considered as being partitioned binomially into $\tilde{R}$ and $\tilde{u}$. At this point, it is useful to introduce the continuous cumulative probability function defined in IBNYR-I:

$$(3.1)\quad \Pi(t\mid\theta) = \frac{1}{T}\int_{(t-T)^+}^{t} F(w\mid\theta)\,dw = \sum_j \pi_j(\theta) = \sum_h \min(I, J+1-h)\,\varphi_h(\theta),$$

with $t = J\Delta$ and $T = I\Delta$, as before Thus, $\Pi(J\Delta\mid\theta)$ is the mass associated with $\tilde{R}$, and each event is unreported with probability $1-\Pi(J\Delta\mid\theta)$ The *total data conditional likelihood* becomes the multinomial·

$$(3.2)\qquad p(\mathcal{D}\mid\theta, n) = \binom{n}{r \quad n-R}\prod_{j=1}^{J}[\pi_j(\theta)]^{r_j}[1-\Pi(J\Delta\mid\theta)]^{n-R}.$$

Let $\tau = \min(T, t) = \Delta\min(I, J)$. Then, given λ, the *total* number of events generated (but not necessarily observed) in $(0, \tau]$ follows the Poisson law with parameter $\lambda\tau$. Setting $u = n-R$ in (3 2) and marginalizing over all values of u, we obtain the final *data likelihood* in terms of the underlying parameters

$$(3.3)\qquad p(\mathcal{D}\mid\lambda, \theta) = \frac{1}{\Pi(r_j!)}\prod_{j=1}^{J}[\pi_j(\theta)]^{r_j}(\lambda\tau)^R e^{-\lambda\tau\Pi(J\Delta\mid\theta)}$$

(The first term is uninformative, and may be dropped) (3.2) should be compared with (4 2) in IBNYR-I (where R was written r), it might, in fact, be argued directly from it. The last term in (3 3) reflects the coupling between $\tilde{\lambda}$ and $\tilde{\theta}$ induced by the data, so that, even if they are *a priori* independent, they will become *a posteriori* dependent

Assuming θ represents a single delay parameter, the traditional point estimates of the parameters, the MLEs $(\hat{\lambda}, \hat{\theta})$, are found from·

$$(3.4)\qquad (\hat{\lambda}\tau)\,\Sigma\pi_j(\hat{\theta}) = R;\qquad \sum\frac{d\pi_j(\hat{\theta})}{d\theta}\left[\frac{r_j}{\pi_j(\hat{\theta})} - \frac{R}{\Sigma\pi_k(\hat{\theta})}\right] = 0.$$

(All sums are over observed intervals only) The second equation can be used to find $\hat{\theta}$ numerically, which is then used in the first equation to give $\hat{\lambda}$ The ML predictor would then be $\hat{u} = \hat{\lambda}\tau - R$.

4. BAYESIAN FORMULATION

As argued in IBNYR-I, we believe that a Bayesian formulation is the natural one for IBNYR problems, since in most applications there will always be rather good prior opinion and relevant experience data about the likely values of $\tilde{\lambda}$ (which will be linked to the number of risk contracts in the portfolio), and about the parameter(s) of the delay distribution (which reflects claim filing delays, administrative flow, adjustment procedures, etc., that are common to all claims in similar lines in each company). No actuary makes estimates in a complete vacuum. The Bayesian approach also has the great advantage of giving a complete *predictive distribution*, which is essential for setting aside portfolio fluctuation reserves.

For consistency with IBNYR-I, we again assume that $\tilde{\lambda}$ and $\tilde{\theta}$ are, *a priori*, independent, with $p(\lambda)$ a *Gamma* (a, b) density. For the rest of this section, we shall leave $f(\ |\theta)$ and $p(\theta)$ in general form, later specializing to exponential delays and another Gamma prior for $\tilde{\theta}$. As in IBNYR-I, these assumptions do not simplify the *joint posterior-to-data density*, $p(\lambda, \theta \mid \mathscr{D})$, because of the coupling term, $\exp[-\lambda\tau\Pi(J\Delta \mid \theta)]$. However, when predicting the number of unreported events, $\tilde{u} = \tilde{n} - R$, we can follow the development in IBNYR-I and show that $\tilde{u}$, given (λ, θ), is Poisson with parameter $\lambda[T - \tau\Pi(t \mid \theta)]$, because of a fortuitous cancellation of the coupling term. Thus, the *predictive density* factors into a product of two *shaping factors*:

$$(4.1) \qquad p(u \mid \mathscr{D}) \propto h_\lambda(u \mid \mathscr{D})\, h_\theta(u \mid \mathscr{D}),$$

with

$$(4.2) \qquad h_\lambda(u \mid \mathscr{D}) = \frac{T^u}{u!} \int \lambda^{R+u} e^{-\lambda T} p(\lambda)\, d\lambda \propto \frac{\Gamma(a+R+u)}{u!} \left[\frac{T}{b+T}\right]^u$$

with a *Gamma* (a, b) prior, and

$$(4.3) \qquad h_\theta(u \mid \mathscr{D}) = \int \prod_{j=1}^{J} [\pi_j(\theta)]^{r_j} \left[1 - \left(\frac{\tau}{T}\right) \Pi(t \mid \theta)\right]^u p(\theta)\, d\theta$$

for Type IIq data, with a similar form for Type Iq. Note that the first shaping factor depends only on R and $p(\lambda)$, while (4.3) depends only on $\boldsymbol{r}$ or $\boldsymbol{s}$ and $p(\theta)$. As in IBNYR-I, we refer to the term involving u in (4.3) as the *kernel*, $K(\theta)$.

Computation of the predictive distribution is most easily accomplished using the recursive form:

$$(4.4) \qquad \frac{p(u+1 \mid \mathscr{D})}{p(u \mid \mathscr{D})} = \left(\frac{a+R+u}{u+1}\right)\left(\frac{T}{b+T}\right)\left(\frac{h_\theta(u+1 \mid \mathscr{D})}{h_\theta(u \mid \mathscr{D})}\right),$$

calculated by starting with $p(0 \mid \mathscr{D}) = 1$, then normalizing when finished. With no data, the marginal (preposterior) predictive density is simply a *Pascal* $(a, T/(b+T))$ density. As in IBNYR-I, (4.4) also provides a Bayesian point estimator, the *predictive mode*, $\hat{u}(\mathcal{D})$, as the smallest integer not less than the value u^* that satisfies:

$$u^*+1 = \left(\frac{a+R+u^*}{b+T}\right) T \left(\frac{h_0(u^*+1 \mid \mathcal{D})}{h_0(u^* \mid \mathcal{D})}\right) \tag{4.5}$$

Note that only the *ratios* of h_0 are needed in (4.4), which means that simple approximations to the integrals will give quite accurate predictive densities (TIERNEY & KADANE, 1986), (KASS, TIERNEY & KADANE, 1988) We now consider how these integrals might be approximated if the delay distribution were exponential.

5 EXPONENTIAL DELAY DISTRIBUTION

Following the example in IBNYR-I, we set $f(w \mid \theta) = \theta \exp(-\theta w)$ $(w \geq 0)$, and recall that

$$\Pi(t \mid \theta) = \left(\frac{\tau}{T}\right)(1-\psi(\theta\tau)\, e^{-\theta(t-\tau)}), \tag{5.1}$$

where the properties of the useful function $\psi(x) = [1-e^{-x}]/x$ were given in that paper

Then, from (2 2), we find

$$\Phi_h = I^{-1}(1-\psi(\theta\Delta)\, e^{-(h-1)\theta\Delta}) \qquad (h = 1, 2, \ldots). \tag{5.2}$$

and the Type Iq probabilities from (2 5) are

$$\varphi_h(\theta) = \begin{cases} I^{-1}[1-\psi(\theta\Delta)] & (h = 1) \\ I^{-1}[\theta\Delta\, \psi^2(\theta\Delta)\, e^{-(h-2)\theta\Delta}] & (h = 2, 3, \ldots) \end{cases} \tag{5.3}$$

The slightly more complicated Type IIq data probabilities are found from (2 8) as:

$$\pi_j(\theta) = \begin{cases} I^{-1}[1-\psi(\theta\Delta)\, e^{-(j-1)\theta\Delta}] & (j = 1, 2, \ldots I) \\ I^{-1}[\theta T\, \psi(\theta\Delta)\, \psi(\theta T)\, e^{-(j-I-1)\theta\Delta}] & (j = I+1, I+2, \ldots) \end{cases}. \tag{5.4}$$

Rewriting h_0 as in IBNYR-I.

$$h_0(u \mid \mathcal{D}) = \int L(\theta \mid \mathcal{D})\, [K(\theta)]^u\, p(\theta)\, d\theta, \tag{5.5}$$

the *epoch data likelihood*, $L(\theta)$, is then expressed for Type Iq data as

$$L(\theta \mid \mathcal{D}) = \prod_{h=1}^{I} [\varphi_h(\theta)]^{N_h} \propto [1-\psi(\theta\Delta)]^{N_1}\, [\theta\Delta\, \psi^2(\theta\Delta)]^{R-N_1}\, e^{-M\theta\Delta}, \tag{5.6}$$

where uninformative constants have been dropped, and M_s is the moment:

$$(5.7)\qquad M_s = \sum_{h=2}^{J} (h-2)\, s_h .$$

In other words, with exponential delays, (s_1, R, M_s) becomes the reduced set of sufficient statistics for Type Iq data. Remember that, with each new value of J, the $I-1$ most recent values of s have to be recomputed from (2.4), otherwise, there is nothing special about the choice of J relative to I.

For Type IIq data, assuming $J > I$:

$$(5.8)\qquad L(\theta \mid \mathcal{D}) = \prod_{j=1}^{J} [\pi_j(\theta)]^{r_j} \propto \prod_{j=1}^{I} [1-\psi(\theta\Delta)\, e^{-(j-1)\theta\Delta}]^{r_j}$$
$$\times\ [\theta T\, \psi(\theta\Delta)\, \psi(\theta T)]^{R_J}\, e^{-M_r \theta\Delta},$$

where uninformative constants have been dropped, and

$$(5.9)\qquad R_J = \sum_{j=I+1}^{J} r_j, \qquad M_r = \sum_{j=I+1}^{J} (j-I-1)\, r_j .$$

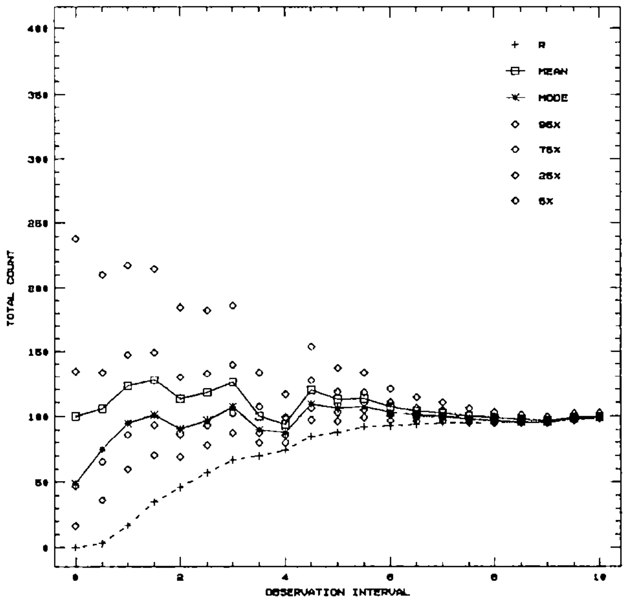
In this case, $(r_1, r_2, \ldots r_I; R_J; M_r)$ become the sufficient statistics If $J \le I$, the product term in (5.8) has an upper limit of J, the terms on the second line are dropped since $R_J = M_r = 0$, and the sufficient statistics revert to $(r_1, r_2, \ldots r_J)$. In contrast to Iq data, once all of the values in $\boldsymbol{r}$ are computed for a given I, they can be used for any J

6 NUMERICAL EXAMPLE AND DATA ANALYSIS

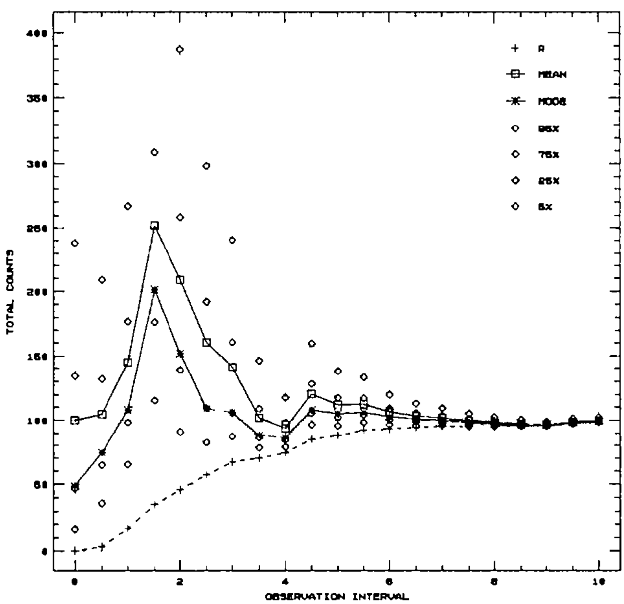
To facilitate comparison with prediction using continuous data, we will use the same basic data and assumptions as in IBNYR-I, namely, that $\tilde{\lambda}$ has a $\mathcal{G}amma$ (2, 0.02) prior density and $T = 1$, so that the no-data (marginal) prediction density is $\mathcal{P}ascal$ $(2, 1\ 02^{-1})$, with mean $\mathcal{E}\{\tilde{n}\} = 100$ events, mode $\hat{n} = 49$, and fractiles $n_{05} = 16\ 5$, $n_{25} = 47.0$, $n_{75} = 134.5$, and $n_{95} = 238\ 1$ The delay is assumed to be exponentially distributed, with a $\mathcal{G}amma$ (4,6) prior density on $\tilde{\theta}$, so that the prior mean delay is $\mathcal{E}\{\tilde{\theta}^{-1}\} = 2.0$ years, with $\mathcal{V}\{\tilde{\theta}^{-1}\} = 8.0$ years^2

For the purposes of simulation, we "stacked the deck" by using the same 100 samples (x_k, y_k) as IBNYR-I, where the x_k were drawn from a uniform distribution over (0,1), and the delays, $w_k = y_k - x_k$, were drawn from an exponential density with true parameter $\theta = 0\ 5$ years^{-1} As shown in Table 1 of IBNYR-I, this gave continuous delay samples from 0 163 to 12.402 years, with a sample average delay of 2 35 years, somewhat larger than the true mean Thus, our experiment assumes accurate but not too precise prior knowledge, so that the behavior below shows primarily the effects of quantization and the two different data types Clearly, with vaguer prior information, we would see a

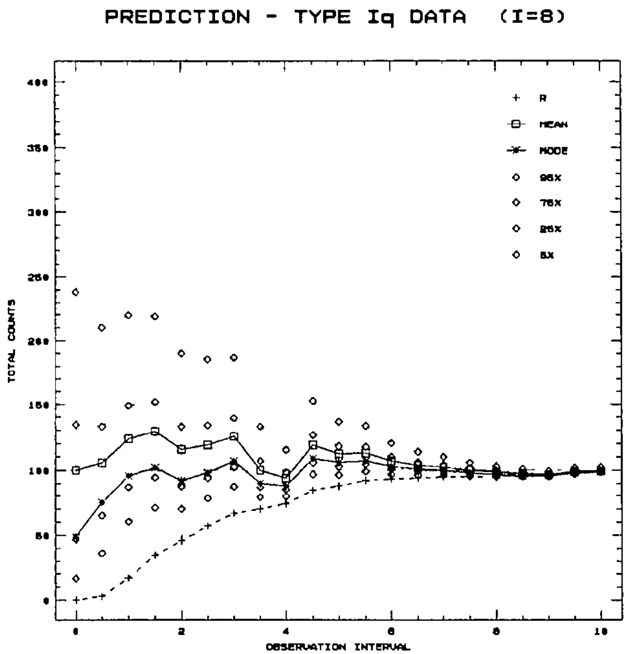
PREDICTION - TYPE Ic DATA

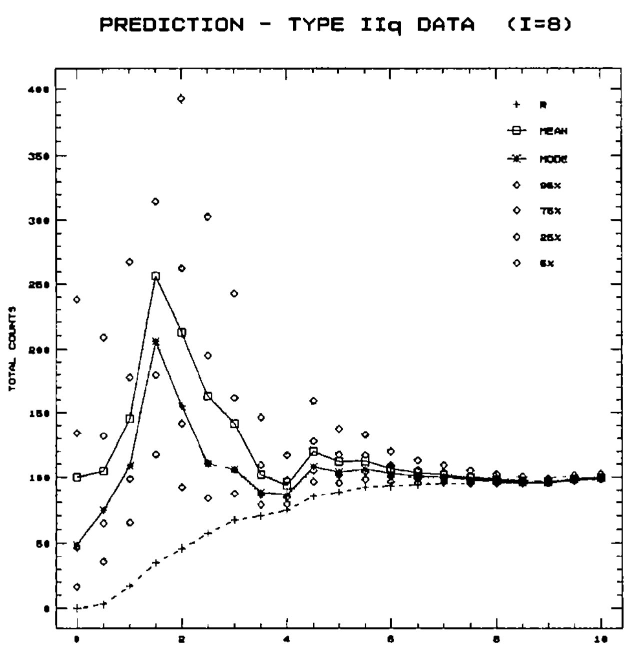


PREDICTION - TYPE IIc DATA



FIGURES 2a & 2b Predictive mean, mode, and fractiles versus t for Types Ic and IIc continuous data





FIGURES 3a & 3b Predictive mean, mode, and fractiles versus t for Types Iq and IIq quantized data ($I = 8$)

further degradation of the predictive power for the smaller values of $t(J)$. Figure 1 shows the individual cell counts for this sample when $\Delta = 0.25$ years ($I = 4$), and $t = 4.0$ years ($J = 16$). The values for the statistics s and r were given above in Section 2.

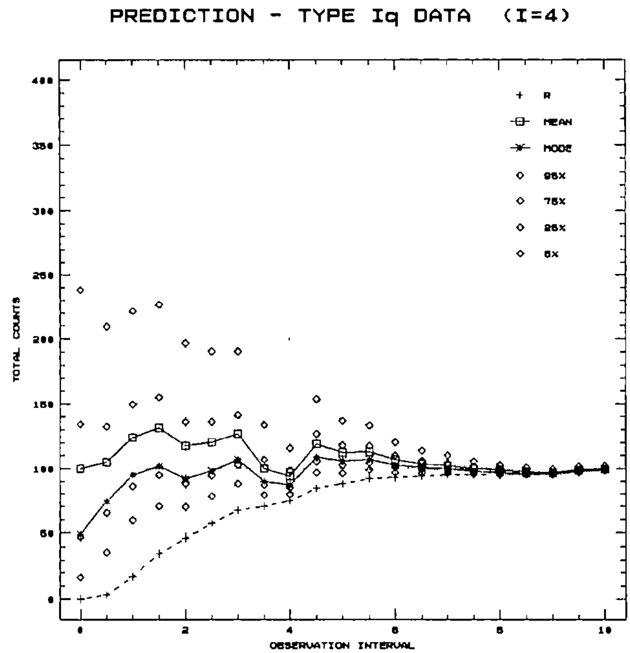
As the effects of quantization are the main interest of this paper, computations were carried out for many different values of I, with $I = 1, 2, 4$, and 8 finally chosen as representative, with complete predictive densities computed for observation intervals $t = 0(0\ 5)\ 10.0$, except when $I = 1$, when only $t = 0(1\ 0)\ 10\ 0$ is possible. Approximations for the shaping factor integral h_0 were computed using the Gammoid method outlined in IBNYR-I, in which a numerical search for the mode, $\hat{\theta}$, of the combination $L(\theta \mid \mathcal{D})\, p(\theta)$ is made, and the unimodal curve then approximated at the mode by a curve of the form $g(\theta) = (A\theta)^G e^{-D\theta}$ Since, to a good approximation, the kernel $K(\theta) \approx e^{-\delta\theta}$ in the neighborhood of this mode, the integral (4.3) can be computed exactly, giving a final recursive relationship like that in (10.1) of IBNYR-I Initially, the mode was chosen from the prior density as $\hat{\theta} = 0.5$, from two to five iterations were then necessary to find the true value of the mode, which ranged from 0 46 to 1.98 in the cases examined. For smaller values of t and I, $p(u \mid \mathcal{D})$ is heavy in the tails, so, to obtain stable means, the recursion (10.1) was carried out over the range [0,1000], and, in a few cases, [0,2000]. As the no-data ($t = 0$) case is known analytically, a total of $2 \times (10 + 3 \times 20) = 140$ complete densities, $p(u \mid \mathcal{D})$, were computed for Figures 3-6 below This task took 5-10 seconds per density on a PC-AT. The densities themselves look much like Figures 5 and 6 in IBNYR-I, and are not shown. But from these, the means, modes, and fractiles shown in the figures below were computed for the total count $n = R + u$

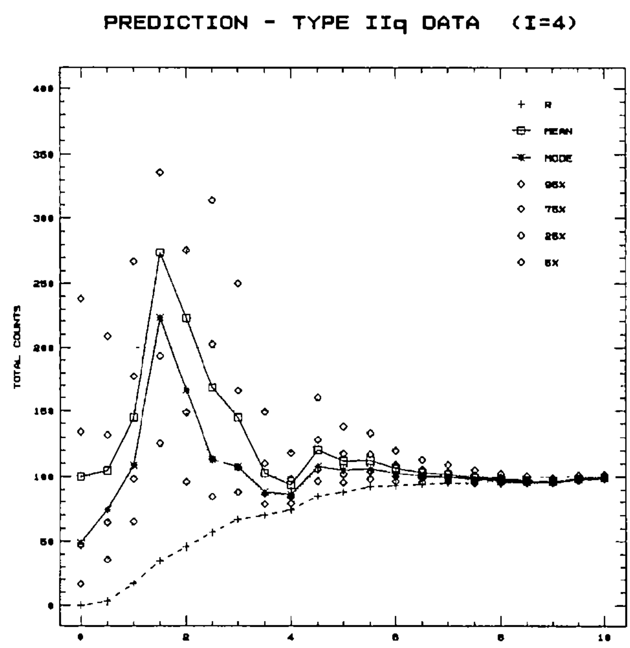
Our standard of comparison will be the continuous data predictions, the results for which are reproduced from IBNYR-I in Figures 2a & 2b; for short, we shall refer to these as the Ic and IIc results, respectively For ease in comparison, we keep the same vertical scale in all plots against the observation interval, $t(J)$.

Figures 3a & 3b show the Types Iq and IIq results for a fine quantization level, $I = 8$ At this level, it is practically impossible to see the effects of discrete reporting, as the only differences are a few percent in the upper fractiles in the interval $1.5 \leq t \leq 2\ 5$

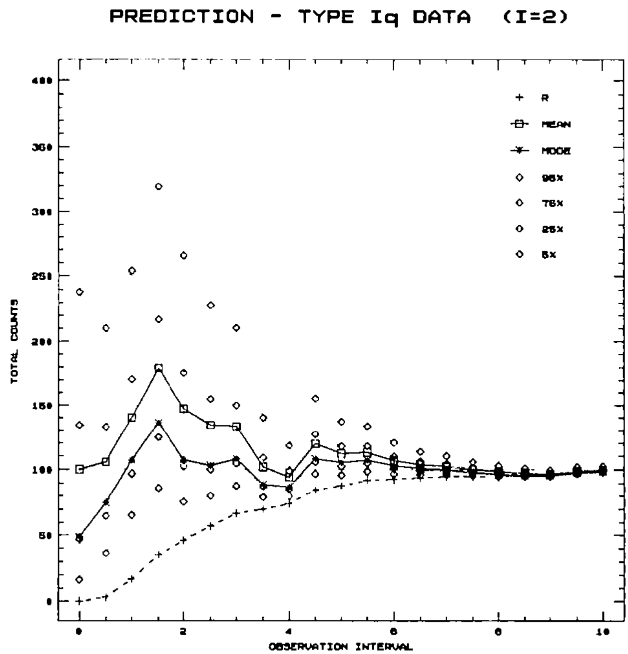
When we coarsen the quantization level to $I = 4$, as shown in Figures 4a & 4b, there begins to be a noticeable increase in the Case Iq upper fractiles and the predictive mean in the interval [1.0, 3 0], but still less than 4% in the worst case. However, the degradation of Type IIq predictions is noticeably worse, with increase in the fractiles, the mean, and the mode in the region [0.5, 3.5], up to 11% in the worst cases It should be remembered that $I = 4$ means that the reporting interval is *one-eighth* the mean delay, which is already more frequent than many implementations encountered in practice

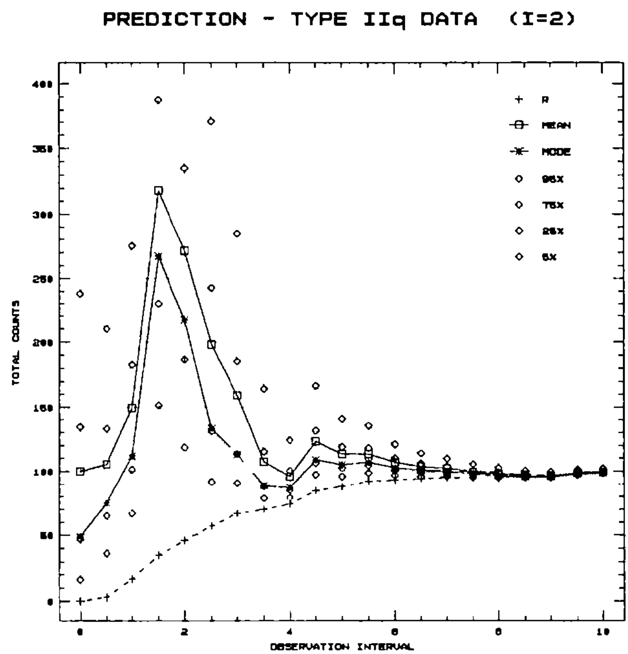
Then, with $I = 2$, Figures 5a & 5b both show the instability in the interval [1 0, 3 5] that before was characteristic of only Type II data. In fact, the Type IIq predictions in the unstable region are now so bad as to be unreliable





FIGURES 4a & 4b Predictive mean, mode, and fractiles versus t for Types Iq and IIq quantized data (I = 4)





FIGURES 5a & 5b Predictive mean, mode, and fractiles versus t for Types Iq and IIq quantized data ($I = 2$)

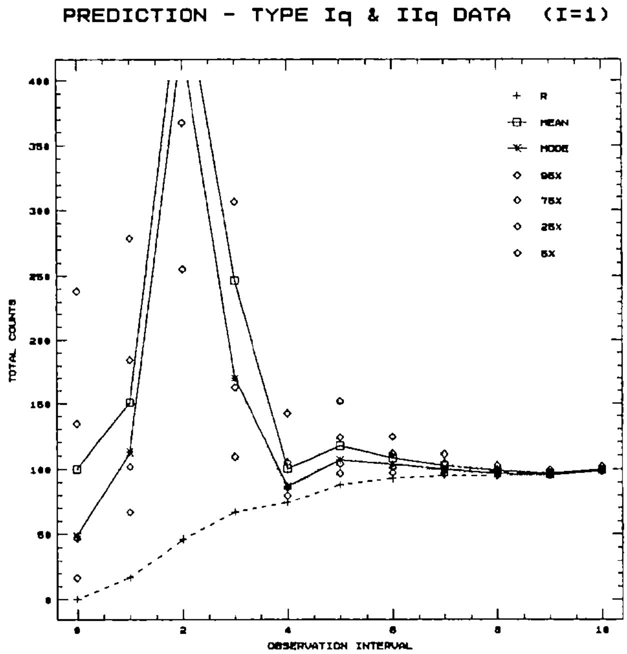


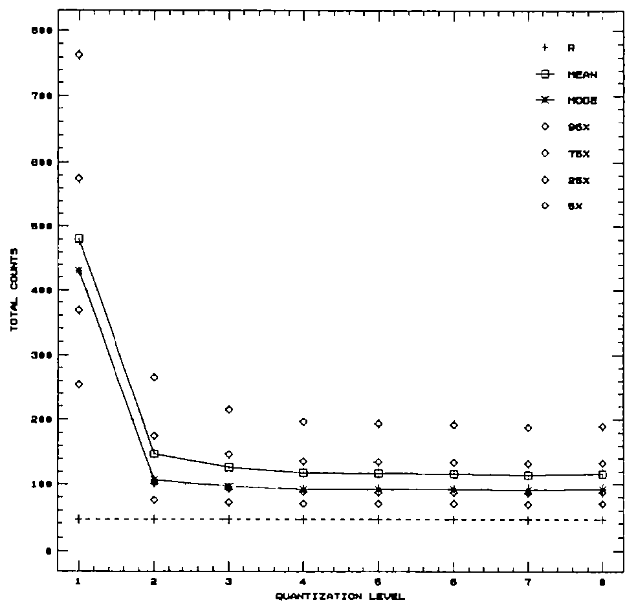
FIGURE 6 Predictive mean, mode, and fractiles versus t for Type Iq = Type IIq quantized data (I = 1)

unless no other estimates are available. Even the region $t \geq 4.0$, which heretofore had given similar results for both types of data because over 74% of the counts were reported, now shows some "bobbling around" due to the changing aggregation of data.

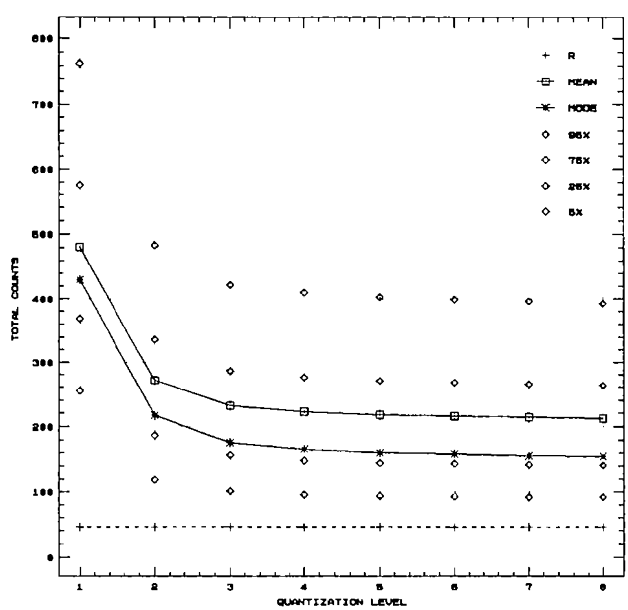
Finally, we have the case $I = 1$ in which Cases Iq and IIq coalesce. To illustrate the extreme degradation in this case, we have chosen to plot the results in Figure 6 on the *same* vertical scale as previous graphs, rather than changing the scale to show all the results. For $t = 2.0$ ($t = 1.5$ cannot be computed), the missing predictive mean count is 481.1, the mode is 430, and the upper fractiles are 575 and 763, respectively! Clearly, the use of a quantization interval that is *one-half* the mean delay is much too coarse when $1.0 \leq t \leq 6.0$. Admittedly, the region above that is reasonable, but that is prediction with at least 93% of the events already reported!

Figures 7a & 7b give a "cross-sectional" impression of the changing level of quantization, in the case for $t = 2.0$, which is in the region of instability with 46% of the events reported. The vertical scale has now been doubled, so that one may now clearly see how bad the cases $I = 1$ and $I = 2$ truly are. In my opinion, one should pick *at least* $I = 4$ in Case Iq and $I = 8$ in Case IIq to get "good" predictions, which means that, given a mean delay of 2.0 years, one must have semi-annual or quarterly data, respectively!

PREDICTION - TYPE Iq DATA (t=2.0)



PREDICTION - TYPE IIq DATA (t=2.0)



FIGURES 7a & 7b Predictive mean, mode, and fractiles versus l for Types Iq and IIq quantized data (t = 2.0)

7. DISCUSSION AND SUMMARY

We should perhaps emphasize once more that the results obtained with changing levels of quantization (for a fixed observation interval) are due solely to changes in Δ and data type upon the epoch data likelihood $L(\theta \mid \mathscr{S})$, in (5.5). This is because the part of the prediction that depends upon λ is unaffected by changing Δ; R reflects *all* of the relevant information we can obtain about the event rate for the purposes of *prediction*. On the other hand, (3.3) shows that the computation of the *joint estimates* of $\tilde{\lambda}$ and $\tilde{\theta}$ will be much more difficult

The effect of quantization upon the epoch data likelihood can be visualized in Figures 8a & 8b, which show this function when $t = 2.0$ for $I = \infty$ (continuous data), 4, 2 and 1, for the two different data types Although the mean and mode shift somewhat as I decreases from ∞ towards 2, the predominant effect is an increased spread in the likelihood These likelihoods are multiplied by the prior density (dotted line), the results approximated by a Gammoid, and then used with the kernel to find the shaping factors $h_0(u \mid \mathscr{S})$, and, from the recursion (4.4), the final predictive density. Note that Type IIq data likelihoods, although converging faster with finer quantization, do not shift the mode as much as Type Iq; since the true value of θ is 0.5 (mode of prior density), this means that Type IIq data will give less accurate predictions. The case $I = 1$ is, well, hopeless.

Keeping in mind the summary observations that were already made in IBNYR-I about the continuous-data prediction problem, the main lessons to be drawn from this paper are·

(1) The introduction of quantized reporting of epochs into the IBNYR model requires no new concepts and only a modest increase in algebra and computational effort.

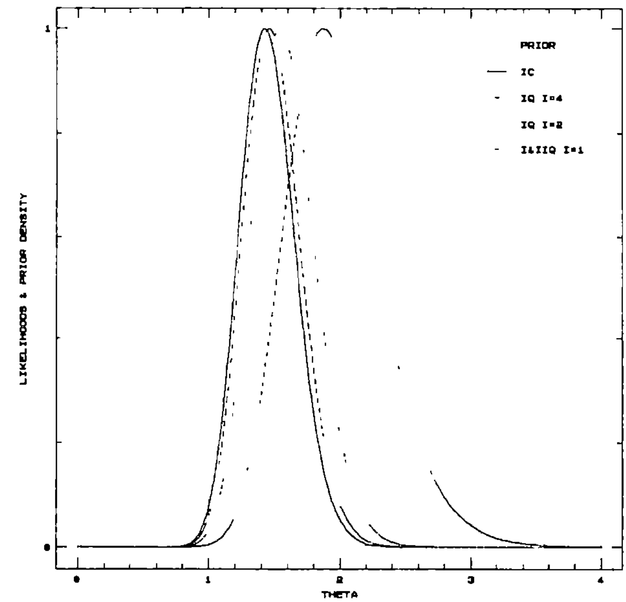
(2) Case IIq data (no occurrence dates reported) continue to give poorer predictions than Case Iq (both occurrence and reporting epochs known) and the predictions degrade more quickly with coarser quantization.

(3) The predictive accuracy of these discrete-time models, in comparison with the continuous case, declines dramatically as Δ increases from, say, one-sixteenth the mean delay to one-quarter the mean delay. A tentative rule-of-thumb seems to be to choose Δ to be at least one-eighth the mean delay with Iq data and one-sixteenth the mean delay with IIq data, if at all possible.

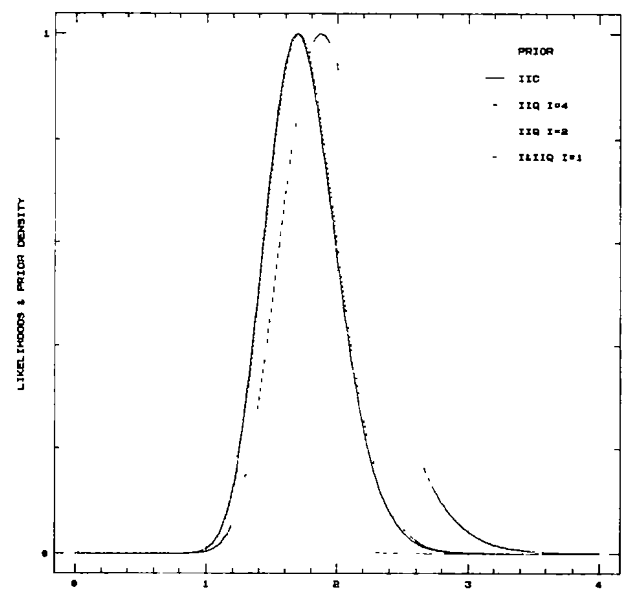
(4) The case $I = 1$ (Δ is one-half the mean delay), while coalescing the two data types and simplifying the sufficient statistics, is so poor as to be unusable in the region of interest.

Admittedly, it is dangerous to extrapolate from one numerical example to practice For instance, one may be able to be much more precise *a priori* about the parameters of the delay distribution; this narrower prior will, to some extent, counteract the imprecise data likelihoods obtained with coarse quantization. And, as always, the final predictive spreads can be greatly reduced if we

EPOCH LIKELIHOODS & PRIOR - TYPE I DATA



EPOCH LIKELIHOODS & PRIOR - TYPE II DATA



FIGURES 8a & 8b Epoch Data Likelihood, $L(\theta|\cdot)$, and Prior Density, $p(\theta)$, versus θ for Types Iq and IIq quantized data (t = 2 0) can provide better prior information about the occurrence rate, perhaps by incorporating the underlying business volume into the model

With this understanding of the potential hazards of quantized reporting, our next paper will consider the question of whether or not *cohort data* from an IBNR traingle can sharpen our estimation of the unknown delay distribution and improve our predictions of the unreported events

I would like to thank M. LIN for her substantial computational and proofing assistance in developing these results Any comments or criticisms on this paper are welcome, as are suggestions for making the basic model more realistic and useful.

ACKNOWLEDGMENT

This research has been partially supported by the Air Force Office of Scientific Research (AFOSR), USAF, under Grant AFOSR-81-0122 with the University of California Reproduction in whole or in part is permitted for any purpose of the United States Government

PARTIAL REFERENCES

JEWELL, W S (1989) Predicting IBNYR events and delays *ASTIN Bulletin* **19**, 25–55

KASS, R E, TIERNEY, L and KADANE, J B (1988) Asymptotics in Bayesian computation In *Bayesian Statistics 3*, J M Bernardo, M H DeGroot and A F M Smith (Eds), Oxford University Press, 261–278

KAHANE, Y (1989) A modern approach to loss reserving in long-tail lines — the case of automobile insurance *Transactions, XXIth ASTIN Colloquium, New York*, 175 197

Reinsurance Association of America (1987) Loss development study *Transactions, XXIth ASTIN Colloquium, New York*, 525–583

SCHNIEPER, R (1989) A pragmatic IBNR method *Transactions, XXIth ASTIN Colloquium, New York*, 405–425

STIERS, D (1989) Applications de méthodes d'evaluation des réserves *Transactions, XXIth ASTIN Colloquium, New York*, 441–463

TIERNEY, L and KADANE, J B (1986) Accurate approximations for posterior moments and marginal densities *J Amer Statist Assoc* **81**, 82-86

WELLER, A O (1989) Generalized Bondy development *Transactions, XXIth ASTIN Colloquium, New York*, 585–609

WILLIAM S. JEWELL
Department of Industrial Engineering & Operations Research, 4173 Etcheverry Hall, University of California, Berkeley, California 94720 USA.

SHORT CONTRIBUTIONS

RUIN PROBABILITY FOR TRANSLATED COMBINATION OF EXPONENTIAL CLAIMS

By Beda Chan

University of Toronto, Canada

Abstract

An alternative expression for the coefficients in the ruin probability for the classical ruin model with translated combination of exponential claims is derived

Keywords

Probability of ruin; translated combination of exponentials

In a compound Poisson claim process with claim amounts distributed as a mixture of exponentials

$$p(x) = \sum_{i=1}^{n} A_i \beta_i e^{-\beta_i x}$$

for $x > 0$ where all $A_i > 0$ and $\sum_{i=1}^{n} A_i = 1$, it is well known that the ruin probability is also a linear combination of exponentials

$$\psi(u) = \sum_{i=1}^{n} C_i e^{-r_i u}$$

where $\{r_1, \ldots, r_n\}$ are solutions to the adjustment coefficient equation

$$(1+\theta)\, p_1 = \frac{M_X(r)-1}{r}$$

and $\{C_1, \ldots, C_n\}$ are determined by the partial fractions of

$$\sum_{i=1}^{n} \frac{C_i r_i}{r_i - r} = \frac{\theta}{1+\theta} \cdot \frac{\dfrac{M_X(r)-1}{r}}{(1+\theta)\, p_1 - \dfrac{M_X(r)-1}{r}} .$$

See Bowers et al. (1986), § 12 6 for details. This result was later extended by Dufresne and Gerber (1989) to the case when the claim distribution is a translated (density function moved by τ to the left) combination of exponentials. (Note that the A_i's need not be positive) They found that the coefficients C_i's are the solution to the system:

$$\sum_{k=1}^{n} \frac{\beta_i}{\beta_i - r_k} C_k = 1, \qquad i = 1, \ldots, n, \tag{1}$$

and gave C_k explicitly. In this note we give an alternative expression for the solution for (1):

$$C_k = \prod_{\substack{i \neq k \\ i=1}}^{n} \frac{r_i}{r_i - r_k} \prod_{i=1}^{n} \frac{\beta_i - r_k}{\beta_i}. \tag{2}$$

To verify (2), consider

$$\sum_{i=1}^{n} \frac{x}{x - r_i} C_i = 1 - \prod_{i=1}^{n} \frac{r_i(x - \beta_i)}{\beta_i(x - r_i)}$$

where the two sides are different expressions for the same rational function of (degree n/degree n) which has simple poles $\{r_1, \ldots, r_n\}$ and takes the value 1 at $x = \beta_1, \ldots, \beta_n$ and the value 0 at $x = 0$. Multiply by $x - r_k$ and let $x = r_k$ to obtain (2).

Two different expressions for C_k, (49) and (54) in DUFRESNE and GERBER (1989), arise naturally when a more detailed problem including the severity of ruin is studied These two expressions can be obtained from summing (9) and (22) in DUFRESNE and GERBER (1988) over j respectively

ACKNOWLEDGEMENT

The author gratefully acknowledges these contributions of colleagues · Professor FUNG-YEE CHAN solved (1) for $n = 2, 3$, and 4 using MAPLE, a language for symbolic computing. Professor HANS U GERBER offered (2) as an exercise. Professor ELIAS S.W. SHIU arranged for a visit to friendly Manitoba during the tenure of Professor GERBER as the inaugural Dr. L.A.H. WARREN, Professor in Actuarial Science, Faculty of Management, the University of Manitoba

REFERENCES

BOWERS, N L JR, GERBER, H U, HICKMAN, J C, JONES, D A and NESBITT, C J (1986) *Actuarial Mathematics* Society of Actuaries, Schaumburg, Illinois

DUFRESNE, F and GERBER, H U (1988) The Probability and Severity of Ruin for Combination of Exponential Claim Amount Distributions and Their Translations *Insurance Mathematics and Economics* **7**, 75–80

DUFRESNE, F and GERBER, H U (1989) Three Methods to Calculate the Probability of Ruin, *ASTIN Bulletin* **19**, 71–90

BEDA CHAN

Department of Statistics, University of Toronto, Toronto, Canada M5S 1A1.

BOOK REVIEWS

W R HEILMANN (1988). *Fundamentals of Risk Theory.* Verlag für Versicherungswirtschaft, Karlsruhe, 288 pages, 36 DM.

This book is essentially an English translation of the book "Grundbegriffe" by the same author. Our readers are therefore referred to the book review of "Grundbegriffe" published in the *ASTIN Bulletin* **18**, vol 1, 115–116.

* * *

IBR. International Bibliography of Reinsurance 9th edition 1989/90 Copyright and edited by. Bayerische Rückversicherung Aktiengesellschaft, Sederanger 4–6, Tucherpark, D-8000 München 22.

The present bibliography which appeared for the first time in 1962 documents some 4'800 titles in reinsurance literature going back to 1912. Earlier books and articles have been included only if of historical value or of perennial fundamental importance.

In the present ninth edition the new publications of 1987 and 1988 have been included. The closing date for inclusion was 31st December 1988.

IBR has remained a handy reference book which as was already mentioned in the preface of the first edition "is intended to serve the purposes of practical insurance by helping those working on a reinsurance problem to turn to account the experiences, considerations and proposed solutions as reported in publications". Because it is compiled simultaneously in two languages—German and English—the IBR can be sure to find a large and international readership.

ERWIN STRAUB

GUIDELINES TO AUTHORS

1. Papers for publication should be sent in quadruplicate to one of the Editors

   Hans Buhlmann,
   Mathematik, ETH-Zentrum,
   CH-8092 Zurich, Switzerland

   D Harry Reid,
   Eagle Star Insurance Company Ltd,
   Eagle Star House, 9 Aldgate High Street,
   London EC3N 1LD, United Kingdom

   or to one of the Co-Editors

   Alois Gisler,
   "Winterthur" Swiss Insurance Company,
   P O Box 357, CH-8401 Winterthur, Switzerland

   David Wilkie
   Messrs R Watson & Sons
   Watson House, London Rd, Reigate, Surrey RH2 9PQ, United Kingdom

   Submission of a paper is held to imply that it contains original unpublished work and is not being submitted for publication elsewhere
   Receipt of the paper will be confirmed and followed by a refereeing process, which will take about three months

2. Manuscripts should be typewritten on one side of the paper, double-spaced with wide margins The basic elements of the journal's style have been agreed by the Editors and Publishers and should be clear from checking a recent issue of *ASTIN BULLETIN* If variations are felt necessary they should be clearly indicated on the manuscript

3. Papers should be written in English or in French. Authors intending to submit longer papers (e g exceeding 30 pages) are advised to consider splitting their contribution into two or more shorter contributions

4. The first page of each paper should start with the title, the name(s) of the author(s), and an abstract of the paper as well as some major keywords An institutional affiliation can be placed between the name(s) of the author(s) and the abstract

5. Footnotes should be avoided as far as possible

6. Upon acceptance of a paper, any figures should be drawn in black ink on white paper in a form suitable for photographic reproduction with lettering of uniform size and sufficiently large to be legible when reduced to the final size

7. References should be arranged alphabetically, and for the same author chronologically Use a, b, c, etc to separate publications of the same author in the same year For journal references give author(s), year, title, journal (in italics, cf point 9), volume (in boldface, cf point 9), and pages. For book references give author(s), year, title (in italics), publisher, and city

   Examples

   BARLOW, R E and PROSCHAN, F (1975) *Mathematical Theory of Reliability and Life Testing* Holt, Rinehart, and Winston, New York
   JEWELL, W S (1975a) Model variations in credibility theory In *Credibility Theory and Applications* (ed P M KAHN), pp 193–244, Academic Press, New York
   JEWELL, W S (1975b) Regularity conditions for exact credibility *ASTIN Bulletin* **8**, 336–341

   References in the text are given by the author's name followed by the year of publication (and possibly a letter) in parentheses

8. The address of at least one of the authors should be typed following the references

Continued overleaf

COMMITTEE OF ASTIN

Guidelines to Authors *continued from inside back cover*

9. Italics (boldface) should be indicated by single (wavy) underlining Mathematical symbols will automatically be set in italics, and need not be underlined unless there is a possibility of misinterpretation
Information helping to avoid misinterpretation may be listed on a separate sheet entitled 'special instructions to the printer' (Example of such an instruction Greek letters are indicated with green and script letters with brown underlining, using double underlining for capitals and single underlining for lower case)

10. Authors will receive from the publisher two sets of page proofs together with the manuscript One corrected set of proofs plus the manuscript should be returned to the publisher within one week Authors may be charged for alterations to the original manuscript

11. Authors will receive 50 offprints free of charge Additional offprints may be ordered when returning corrected proofs A scale of charges will be enclosed when the proofs are sent out