INDIVIDUAL CLAIMS SIMULATION MACHINE: DESCRIPTION OF THE R PACKAGE

Andrea Gabrielli

Mario V. Wüthrich

Version V1, February 26, 2018

Abstract

The R files Functions.V1 and Simulation.Machine.V1 serve at simulating individual claims cash flows of non-life insurance claims for a synthetic insurance portfolio.

Description

The R files Functions.V1 and Simulation.Machine.V1 serve at simulating individual claims cash flows of non-life insurance claims for a synthetic insurance portfolio. The construction of the simulation machine is based on a number of different neural networks that have been calibrated to real insurance data coming from accident insurance business. It considers for each individual claim the 6 features LoB, cc, AY, AQ, age and inj_part described below. Based on these 6 features, it simulates a reporting delay for each individual claim and cash flows (thereafter) of a maximal settlement delay of 11 years (including the reporting delay). These cash flows also consider potential recoveries (with a negative sign). Finally, an indicator process is simulated that shows whether the claim is open or closed at the end of each accounting year. This indicator process also considers possible re-openings. For a detailed description of the statistical model used we refer to [1]. Moreover, Functions.V1 also includes a function that allows for simulating a synthetic claims portfolio. The file Functions.V1 contains all the necessary R functions and the file Simulation.Machine.V1 demonstrates their use.

Simulation.Machine(features, npb, seed1, std1, std2)

Description

This is the main function that generates the individual claims cash flows based on the given portfolio of claims features features.

Arguments

features

This is a matrix that contains the claims portfolio. It consists of one row per individual claim and of d=7 columns having the following structure: (ClNr, LoB, cc, AY, AQ, age, inj_part). The variable ClNr is the claim number, this is a unique numerical identifier for each individual claim. LoB is the line of business which should be of factor type (categorical) and may take the labels $1, \ldots, 4$. cc is the claims code that should be of factor type (categorical) and may take the labels $1, \ldots, 53$ (with possible gaps). AY is the year of claims occurrence (accident year) that should be of integer type and may take the values $1994, \ldots, 2005$. AQ is the quarter of claims occurrence (accident quarter) that should be of integer type and may take the values $1, \ldots, 4$. age is the age of the injured that should be of integer type and may take the values $15, \ldots, 70$. inj_part is the body part injured that should be of factor type (categorical) and may take the labels $1, \ldots, 99$ (with gaps).

npb For parallel computing the maximal size of the blocks is specified.

seed1 The seed for the simulation is initialized.

std1 Standard deviation parameter in the log-normal distribution for the total claim amounts simulation.

std2 Standard deviation parameter in the log-normal distribution for the totual recovery payments simulation.

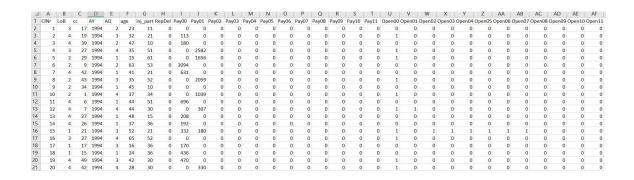
Output

features The output contains the input features (ClNr, LoB, cc, AY, AQ, age, inj_part).

RepDel Reporting delay of the claims (difference between claims occurrence and claims reporting) being in 0, ..., 11.

Pay00,... Cash flows paid in development years (delays) j = 0,...,11.

Open00,... Indicator showing whether the claim was open (=1) or closed (=0) at the end of development year $j=0,\ldots,11$.



Feature.Generation(V, LoB.dist, inflation, seed1)

Description

This function generates a synthetic portfolio of claims features features.

Arguments

V Totally expected number of claims (used for Poissonian claims counts).

LoB.dist Probabilities specifying how the claims counts are allocated to the 4 lines of

business.

inflation Growth parameter for each line of business acting on the claims counts.

seed1 The seed for the simulation is initialized.

Output

The output contains a portofolio with features (ClNr, LoB, cc, AY, AQ, age, inj_part).

1	Α	В	C	D	E	F	G
1	ClNr	LoB	СС	AY	AQ	age	inj_part
2	1	3	17	1994	2	23	11
3	2	4	19	1994	3	32	21
4	3	4	39	1994	2	47	10
5	4	3	27	1994	4	35	51
6	5	2	29	1994	1	15	61
7	6	2	9	1994	2	63	53
8	7	4	42	1994	1	41	21
9	8	2	43	1994	3	35	52
10	9	2	34	1994	1	45	10
11	10	2	1	1994	4	37	34
12	11	4	6	1994	1	44	51

Acknowledgment

Our greatest thanks go to Suva, Dr. Peter Blum and Dr. Olivier Steiger for providing data and for their immense support.

References

[1] Gabrielli, A., Wüthrich, M.V. (2018). Individual claims history simulation machine. SSRN Manuscript ID 3130560.