Insurance Problems Obtained Using the Software Package Crac 2.0

Virginia ATANASIU
Department of Mathematics, Academy of Economic Studies
e-mail: virginia_atanasiu@yahoo.com

The paper analyses some practical problems of health insurance that can be solved by means of credibility theory. All numerical results in this paper were obtained using the software package CRAC 2.0, which uses Jewell's hierarchical model. Examples of the output of the program will be given and discussed. This will give more insight and understanding of the theoretical aspects and will point the way to the practical possibilities of the credibility models.

Mathematics Subject Classification: 62P05.
***Keywords**: the risk premium per employee insured, the sectorial premium, the global premium, an example in health insurance.*

Introduction

Here we consider applications of credibility theory dealing with real life situations, and implemented on real insurance portfolios. Though more examples could be given, we limit ourselves to the introduction of a problem of health insurance. In this example we try to demonstrate what kind of data is needed to apply credibility theory.

Starting from a real portfolio of contracts a risk premium will be calculated on three different levels: the overall portfolio level, the in-between or sector level (where sectors have to be chosen and defined by the user) and the individual policy-level.

1. General information

Description of health insurance

In the example of health-insurance considered here, a firm insures all its employees as a group for medical costs. The clients of the insurance company are firms, not individual employees.

It has been decided that the subdivision of the portfolio of these contracts will be based on the following criteria (the values between brackets give the corresponding sector number in the CRAC software).

- kind of activities of the firm: metal (1), chemical (2), agricultural (3), kind 4 (4), kind 5 (5), kind 6 (6) and food (7);
- geographical positions of the firm: North (1), Central (2), South (3);
- number of employees of the firm: we defined 7 different groups from small (1) to immense (7).

Notice that the portfolio could be split up according to three criteria: the first and the third have values ranging from 1 to 7, the second has values ranging from 1 to 3. By combining all criteria we obtain a subdivision into 147 = 7·3·7 sectors.

1.2 Description of the problem

In this example we want to calculate a risk premium per employee insured. We are especially interested in:

- a risk premium per sector, where sectors are defined according to the criterion "number of employees";
- conceivably an individual premium for more important firms.

To solve this practical problem we make the following choices for (1):

- numerator = amount paid during observation period – t (t = 1,2,3) by the insurance company;
- deductible = 0; we are not interested in calculating the effect of a deductible on the risk premium;
- scaling factor = 1;
- denominator = number of employees within the insured company, where:

$$X_{pjr} = \frac{\left((numerator)_{pjr} - deductible\right)_+}{factor \cdot (deno\min ator)_{pjr}} \quad (1),$$

with: q_{+}=max{0,q}, p denotes the sector this policy belongs to, j denotes the contract and r denotes the observation period. For the numerator we usually take claim payments, while the denominator expresses some measure of exposure like premium volume or number of participants.

2. Practical solution

Example of the input data

In the table below, an example is given of how the input data might look. The first record of this file contains information on policy Z 1024. The three following fields, each of length 2, have the values 030302, indicating that this policy covers an agricultural firm (03) in the Southern part (03) of the country; on the basis of the number of employees, the firm is classified as „size 2" (02).

Next, we see that there is only one observation period available, namely period -3. In this period a amount of 00092357 has been paid; the number of employees insured equals 000045 and the weight has been chosen equal to this number of employees. The interpretation of the other records is straightforward.

Z1024	03	03	02	1	0	00000000	000000	000000	0	00000000
Z1028	06	01	03	3	1	00194514	000547	000547	1	00095224
Z1185	06	01	05	3	1	00039453	001036	001036	1	00141023
Z1188	06	03	05	3	1	00030216	000621	000621	1	00258241

000000	000000	1	00092357	000045	000045
000547	000547	1	00109245	000547	000547
001036	001036	1	00363791	001036	001036
000621	000621	1	00453345	000621	000621

The input data has to be checked first for various errors, such as denominators that are zero, sectors that do not exist, and so on. In our demonstration run 3 out of 703 policies had to be rejected in advance.

Construction of sectors

We could make a subdivision of this portfolio in seven different ways:

1.based on kind of activities: 7 sectors;
2.based on geographical position: 3 sectors;
3.based on a number of employees: 7 sectors;
4.based on a combination of (1.) + (2.):7 × 3 = 21 sectors;
5.based on a combination of (1.) + (3.):7 × 7 = 49 sectors;
6.based on a combination of (2.) + (3.):3 × 7 = 21 sectors;
7.based on a combination of (1.) + (2.) + (3.):7 × 3 × 7 = 147 sectors.

In this application we will use subdivision 3. In the following section we will comment on the credibility results generated by the software for this particular case and subdivision.

Credibility calculations

We give an overview of the method used to obtain the results that follow. In order to facilitate an easy interpretation of the output, we will first discuss some notations used.

We use Jewell's hierarchical model that gives us the possibility to get results on three different levels: global portfolio level, sector level and individual policy level. We start the calculations, beginning with the (transformed) observation X_{pjr} for period r, for policy j in sector p, and calculate successively and iteratively the following weighted averages:

$$X_{pjw} = \sum_{r=1}^{t} X_{pjr} \cdot w_{pjr} / \sum_{r=1}^{t} w_{pjr}$$

$$X_{pzw} = \sum_{j=1}^{k_p} X_{pjw} \cdot z_{pj} / \sum_{j=1}^{k_p} z_{pj}$$

$$X_{zzw} = \sum_{p=1}^{P} X_{pzw} \cdot z_{p} / \sum_{p=1}^{P} z_{p}$$

X_{pjw} measures the average individual experience, weighted by the volumes w_{pjr}, X_{pzw} is the credibility-weighted average sectorial experience, and X_{zzw} is the global experience, again weighted by credibility weights. The determination of the (optimal) credibility weights z_{pj} (of contract j in sector p) and z_p (of sector p) is the most essential part of the calculations outlined above. Finally, the optimal credibility estimators are given as:

global result: $m = X_{zzw}$

sectorial result: $N_p^a = z_p X_{pzw} + (1 - z_p) X_{zzw}$

individual result:

$M_{pj}^a = z_{pj} X_{pjw} + (1 - z_{pj}) X_{pzw}$

2.4. Discussion of the computer output

Next we will discuss the results as they were obtained using the credibility software package CRAC 2.0.. We can distinguish three different parts in these calculations. We will give the computer output and a brief discussion of each of those separately in the sequel.

Results on the level of individual records

Let TAU and Q be technical notations for:

$$\text{TAU} = \sum_{p,j} (t_{pj} - 1)_+ \; ; \; \text{Q} = \sum_p (k_p - 1)_+ ,$$

where t_{pj} is the number of observation periods available for contract j in sector p, and k_p is the number of policies in sector p. The following output results:

CRAC 2.0.

Demonstration

Amount deduced: 0 Scaling factor: 1,0000

Number of sectors: 7 (SIZE OF FIRM)

First step: Calculation time: 0 min. 14 sec.

SUMMARY OF CREDIBILITY RESULTS

S2: 15874241,63527061

TAU: 691

Q: 693

PART 1: ITERATION 1.

Sector 1	Records:	115	Z_{psigma}: 109,25	X_{pzw}: 1,124.9633
Sector 2	Records:	228	Z_{psigma}: 216,60	X_{pzw}: 600,6631
Sector 3	Records:	281	Z_{psigma}: 266,95	X_{pzw}: 322,2896
Sector 4	Records:	42	Z_{psigma}: 39,90	X_{pzw}: 391,0589
Sector 5	Records:	22	Z_{psigma}: 20,90	X_{pzw}: 218,9946
Sector 6	Records:	10	Z_{psigma}: 9,50	X_{pzw}: 252,5414
Sector 7	Records:	2	Z_{psigma}: 1,90	X_{pzw}: 197,2705

$\hat{a}$: 420149,8726747094

PART 1: ITERATION 2.

Sector 1	Records:	115	Z_{psigma}: 45,07	X_{pzw}: 962,1075
Sector 2	Records:	228	Z_{psigma}: 145,25	X_{pzw}: 532,6794
Sector 3	Records:	281	Z_{psigma}: 239,68	X_{pzw}: 270,2506
Sector 4	Records:	42	Z_{psigma}: 34,54	X_{pzw}: 244,6664
Sector 5	Records:	22	Z_{psigma}: 20,07	X_{pzw}: 218,3994
Sector 6	Records:	10	Z_{psigma}: 8,03	X_{pzw}: 217,0130
Sector 7	Records:	2	Z_{psigma}: 1,99	X_{pzw}: 197,2746

$\hat{a}$: 183411,9473936435

……………………….

PART 1: ITERATION 5.

Sector 1	Records:	115	Z_{psigma}: 16,62	X_{pzw}: 846,0575

Sector 2	Records:	228	Z_{psigma}: 70,33	X_{pzw} : 487,7522
Sector 3	Records:	281	Z_{psigma}: 108,34	X_{pzw} : 241,8968
Sector 4	Records:	42	Z_{psigma}: 27,89	X_{pzw} : 195,9236
Sector 5	Records:	22	Z_{psigma}: 16,99	X_{pzw} : 219,5249
Sector 6	Records:	10	Z_{psigma}: 0,55	X_{pzw} : 195,0376
Sector 7	Records:	2	Z_{psigma}: 1,97	X_{pzw} : 197,2900

$\hat{a}$: 71851,24302918158.

PART 1: ITERATION 6.

Sector 1	Records:	115	Z_{psigma}: 14,32	X_{pzw} : 831,8961
Sector 2	Records:	228	Z_{psigma}: 62,10	X_{pzw} : 483,1644
Sector 3	Records:	281	Z_{psigma}: 170,54	X_{pzw} : 238,8752
Sector 4	Records:	42	Z_{psigma}: 26,94	X_{pzw} : 191,9623
Sector 5	Records:	22	Z_{psigma}: 16,48	X_{pzw} : 219,8989
Sector 6	Records:	10	Z_{psigma}: 6,34	X_{pzw} : 192,7394
Sector 7	Records:	2	Z_{psigma}: 1,97	X_{pzw} : 197,2942

$\hat{a}$: 62380,37178764759

……………………

PART 1: ITERATION 9.

Sector 1	Records:	115	Z_{psigma}: 11,13	X_{pzw} : 809,8773
Sector 2	Records:	228	Z_{psigma}: 49,91	X_{pzw} : 476,3438
Sector 3	Records:	281	Z_{psigma}: 153,76	X_{pzw} : 234,2498
Sector 4	Records:	42	Z_{psigma}: 25,32	X_{pzw} : 186,4746
Sector 5	Records:	22	Z_{psigma}: 15,60	X_{pzw} : 220,6503
Sector 6	Records:	10	Z_{psigma}: 5,99	X_{pzw} : 189,2427
Sector 7	Records:	2	Z_{psigma}: 1,95	X_{pzw} : 197,3030

$\hat{a}$: 49032,59194736778

PART 1: ITERATION 10.

Sector 1	Records:	115	Z_{psigma}: 10,62	X_{pzw} : 806,0427
Sector 2	Records:	228	Z_{psigma}: 47,87	X_{pzw} : 475,1882
Sector 3	Records:	281	Z_{psigma}: 150,62	X_{pzw} : 233,4476
Sector 4	Records:	42	Z_{psigma}: 25,01	X_{pzw} : 185,5961
Sector 5	Records:	22	Z_{psigma}: 15,44	X_{pzw} : 220,8065
Sector 6	Records:	10	Z_{psigma}: 5,92	X_{pzw} : 188,6382
Sector 7	Records:	2	Z_{psigma}: 1,95	X_{pzw} : 197,3049

$\hat{a}$: 46864,76682185131

In this example we asked ten iterations (those with number 3, 4, 7 and 8 are not listed). For each of these iterations an overview is given per sector of the number of records, belonging to that sector. Furthermore the quantities X_{pzw} and Z$_{psigma}$ $\left(\sum_j z_{pj}\right)$ are calculated. In order to start the iterative calculation procedure we choose an initial value of 0,95 for the z_{pj}'s. So, for the first iteration, we simply have Z$_{psigma}$ equal to the number of contracts, multiplied with 0,95(115×0,95=109,25). From these initial z_{pj}-values, we calculate the X_{pzw}'s and the corresponding $\hat{a}$-value. This $\hat{a}$-value can then be used in the second iteration to calculate more accurate z_{pj}-values, and so on. Note that we choose a rather high initial value for the z_{pj}, hoping that the numerical process converges to the right value $\hat{a}$>0, if it exists, instead of to $\hat{a}$= 0, which is also a solution of the equation to be solved. The same applies in case $\hat{b}$ and z_p are computed. By comparing the results of the different iterations we see that the Z$_{psigma}$, X_{pzw} and $\hat{a}$-values converge to limit values that will be used as a basis to perform the calculations in the second part. By dividing each of the X_{pzw}-results (p = 1, 2,…,7) of the last iteration by the number of records in the sector, we get some average individual credibility weight. For sector 1 (the sector that contains the small companies) this average credibility weight equals 10,62/115 or 9,23%, whereas for sector 6 we get an average weight of 59,2%; for the sector with the biggest companies this weight equals 97,5%. It is clear that the credibility weights in this sector are very high, because the individual firms are very important and have many employees insured, so that the claims experience of these policies becomes very reliable. At this point we have sufficient information to calculate the individual credibility premium M^a_{pj}. In the next part of this section, the credibility weights z_p for the sectors will be determined. By taking a mixture of the global portfolio result and the sectorial X_{pzw}-values, using the z_p weights, we will find the final sectorial N^a_p.

Results on the sectorial and portfolio level:

The CRAC-output for the demonstration run looks as follows:

PART 2: ITERATION 1.

Sector 1	Records: 115	z_p : 0,949999990
Sector 2	Records: 228	z_p : 0,949999990
Sector 3	Records: 281	z_p : 0,949999990
Sector 4	Records: 42	z_p : 0,949999990
Sector 5	Records: 22	z_p : 0,949999990
Sector 6	Records: 10	z_p : 0,949999990
Sector 7	Records: 2	z_p : 0,949999990

m: 329,575; Z$_{sigma}$: 6,650; b: 51,835829

PART 2: ITERATION 2.

Sector 1	Records: 115	z_p : 0,921514810
Sector 2	Records: 228	z_p : 0,981462420
Sector 3	Records: 281	z_p : 0,994033460
Sector 4	Records: 42	z_p : 0,965109590
Sector 5	Records: 22	z_p : 0,944670740
Sector 6	Records: 10	z_p : 0,867591500

Sector 7 Records: 2 z_p : 0,683140400
m: 334,623; Z_{sigma}: 6,358; b: 49,900782
…………………………………………
PART 2: ITERATION 5.

Sector 1	Records: 115	z_p : 0,918462630
Sector 2	Records: 228	z_p : 0,980692680
Sector 3	Records: 281	z_p : 0,993782400
Sector 4	Records: 42	z_p : 0,963685990
Sector 5	Records: 22	z_p : 0,942462860
Sector 6	Records: 10	z_p : 0,862754700
Sector 7	Records: 2	z_p : 0,674096700

m: 334,763; Z_{sigma}: 6,336; b: 49,729110
PART 2: ITERATION 6.

Sector 1	Records: 115	z_p : 0,918460970
Sector 2	Records: 228	z_p : 0,980692270
Sector 3	Records: 281	z_p : 0,993782220
Sector 4	Records: 42	z_p : 0,963685150
Sector 5	Records: 22	z_p : 0,942461670
Sector 6	Records: 10	z_p : 0,862752020
Sector 7	Records: 2	z_p : 0,674091760

m: 334,763; Z_{sigma}: 6,336; b: 49,729015

We attained the relative precision of 0,00000100. Again we start with an initial value for the corresponding m and b-value. We use the notation Z_{sigma} for the sum of the z_p values, so that for the first iteration we get $Z_{sigma} = 0{,}95 \times 7 = 6{,}65$. The initial weights converge rather fast to limit values in the sixth iteration. The software also provides an automatic test on the relative precision that has been reached. When the accuracy of the results is sufficient, the iterative procedure is ended before the maximum number of iterations is performed. Note that the quantity m or X_{zzw}, giving the global portfolio result according to credibility theory, is also calculated in this second section. The final m-value is 334,763, which means that, in case all the firms insured would pay the same premium, this risk premium per employee insured should be 334,763. The sectorial premium N_p^a can be seen as a differentiation of this global premium according to the various sectors. It is given in the final part of the output.

Final results:
The output of this third part looks as follows:

PART 3: FINAL RESULTS

S1: SMALL	115	z	0,91846N	767,6150
S2: SIZE 2	228	z	0,98069N	472,4769
S3: SIZE 3	281	z	0,99378N	234,0776
S4: SIZE 4	42	z	0,96369N	191,0130
S5: BIG	22	z	0,94246N	227,3634
S6: VERY BIG	10	z	0,86275N	208,6936
S7: IMMENSE	2	z	0,67409N	242,1037

PART 2: Calculation time: 0 min. 8 sec.

Total cumulated calculation time: 1 min. 2 sec.
After 6 iterations, we obtained z_7 = 0,67409 in the previous section. This result also appears in the table above. The sectorial result is then:
$N_7 = z_7 X_{7zw} + (1 - z_7) X_{zzw} = 0,67409 \times 197,3049 + (1 - 0,67409) \times 334,763 = 242,1037$

This means that, for determining the credibility premium for sector 7, we attach for about two thirds importance to the claims experience of the sector, whereas for the rest, the total portfolio claims experience is taken into account. The other final results can be calculated in a similar way. From these results we learn that instead of one unique risk premium of 334,763, we can use a sectorially differentiated tariff structure, with risk premiums ranging from 191 to 768.

Remark

For sector 1 to 4, the risk premium per employee insured decreases as the number of employees within the firm increases. For sectors 5 to 7 there is a slight increase of the risk premium. A possible explanation for this phenomenon could be the following: especially in the small companies (one to five persons) health insurance policies are sometimes used (instead of unemployment periods) to overcome periods with low economical activities. Bigger companies (e.g. sector 4) are not confronted with this problem of fluctuations in their business volume in the same way; they have a more regular business and the manager of the firm is still able to control all his employees, so that these companies can benefit from an attractive risk premium. When companies are still growing (beyond the size of sector 4), the problem of controlling the personnel becomes critical: employees easily get ill because there is nobody who really cares whether they are present or not and because they know that there are enough other employees able to do their work. This could explain why the risk premium for these firms increases.

Conclusions

This example shows that credibility theory is really a useful tool-perhaps the only existing tool-for such insurance applications. The fact that it is based on complicated mathematics, involving conditional expectations, needs not bother the user more than it does when he applies statistical tools like SAS, GLIM, discriminant analysis, and scoring models. These techniques can be applied by anybody on his own field of endeavor, be it economics, medicine, or insurance. We give a rather explicit description of the input data for program CRAC 2.0. used, only to show that in practical situations there will always be enough data to apply credibility theory to a real insurance. The point we want to emphasize is that practical application of credibility theory is feasible nowadays using appropriate software.

References

[1]. *De Vylder, F.* & *Goovaerts, M.J.*, **Semilinear credibility with several approximating functions**, Insurance: Mathematics and Economics, 4, 155- 162. (Zbl. No. 0167-6687), 1985.

[2].*Gerber, H.U.* & *Jones, D.A.*, **Credibility formulas of the updating type**, in “Credibility theory and application”, Proceedings of the Berkeley Actuarial Research Conference on credibility, Academic Press, New York, 89-109, 1975.

[3]. *Gerber, H.U.*, **Credibility for Esscher premiums**, Mitteilungen der VSVM, 80, 307-312, 1980.

[4]. *Goovaerts M. J., Kaas R., Van Heerwaarden A., E., Bauwelincks T.*, **Insurance Series**, **Volume 3, Effective Actuarial Methods**, Elsevier Science Publishers B.V., 187-211, 1990.

[5]. *Pentikäinen T.*, Daykin C. D., Pesonen M., **Practical Risk Theory for Actuaries**, Chapman & Hall, 1993.

[6]. Sundt B., **An Introduction to Non-Life Insurance Mathematics**, volume of the “Mannheim Series”, 22-54, 1984.