Using Mixed Poisson Distributions in connection with Bonus-Malus Systems[1]

J.F. Walhin*• and J. Paris*

* *Institut de Statistique, Université Catholique de Louvain, Belgium*

• *Le Mans Assurances, Belgique*

Abstract

For the construction of bonus-malus systems, we propose to show how to apply, thanks to simple mathematics, a parametric method encompassing those encountered in the litterature. We also compare this parametric method with a non-parametric one that has not yet been used in the actuarial litterature and that however permits a simple formulation of the stationary and transition probabilities in a portfolio whenever we have the intention to construct a bonus-malus system with finite number of classes.

Keywords

Mixed Poisson distribution, non-parametric estimation, Hofmann's distribution, Bayes theorem, bonus-malus systems, stationary distribution.

1. Introduction

The distribution of the number of car accidents in an automobile portfolio is known to be well fitted by mixed Poisson distributions. Let us assume that the number of car accidents is Poisson distributed for each risk in the portfolio. Of course the portfolio is inhomogeneous and the frequency of the risks differs from each other. We assume it follows a random variable. If we define $N(t)$ as the number of claims in $(0, t]$ with $\Pi(n, t) = \mathsf{P}[N(t) = n]$ = the probability that a risk causes n accidents in t years, we have

$$\Pi(n, t|\Lambda) = e^{-\Lambda t}\frac{(\Lambda t)^n}{n!} \tag{1}$$

$$\Pi(n, t) = \int_0^\infty e^{-\lambda t}\frac{(\lambda t)^n}{n!}dU(\lambda) \tag{2}$$

where Λ is a random variable with cumulative density function $U(\lambda)$.
Lemaire (1985) used a Gamma distribution for Λ. This implies that $N(t)$ follows a Negative Binomial distribution. In this case, the construction of a bonus malus system is very easy.

[1]This paper has been presented at the XXVIIIth ASTIN Colloquium, Cairns 10-13 August

Tremblay (1992) used the Inverse Gaussian distribution for Λ. This implies that $N(t)$ follows a Poisson Inverse Gaussian distribution. The construction of a bonus-malus system seems to be very complicated, needing for example the use of modified Bessel functions.
In this article, we use a more general parametric distribution and a non-parametric distribution to fit an automobile portfolio. The construction of the bonus malus system is easily done using the Bayes theorem and the form of the mixed Poisson distribution.
The following properties will be used in the sequel :

$$\Psi_{N(t)}(s) = \Pi(0, t - ts) \tag{3}$$

$$\Psi_{\Lambda}(s) = \Pi(0, -s) \tag{4}$$

where $\Psi_X(s) = \mathrm{E}[e^{sX}]$ denotes the moment generating function of X.
For comparison purpose, we work with a portfolio published by Bühlmann (1970) and used by Lemaire (1985) and Tremblay (1992).

	Observed	NB fit	PIG fit
0	103704	103781,72	103710,04
1	14075	13892,03	14054,65
2	1766	1882,63	1784,91
3	255	256,17	254,49
4	45	34,93	40,42
5	6	4,77	6,94
6	2	0,65	1,26

Table 1: Reference portfolio

We will use the following notations :

q	number of classes for which the observation is different from 0
N	maximum number of claims per risk

2. Parametric estimation

For the parametric case, we use a three parameters distribution encompassing the Negative Binomial and the Poisson Inverse Gaussian distributions. This distribution is due to Hofmann (1955) and has been discussed by Kestemont and Paris (1985).
The Hofmann's distribution is defined as follows

$$\begin{aligned}
\Pi(0,t) &= e^{-\theta(t)} \\
\Pi(n,t) &= (-1)^n \frac{t^n}{n!} \Pi^{(n)}(0,t) \\
\theta'(t) &= \frac{p}{(1+ct)^a} \\
\theta(0) &= 0
\end{aligned}$$

By integration, we have

$$\begin{aligned}
\theta(t) &= pt \quad \text{if } a = 0 \\
\theta(t) &= \frac{p}{c}\ln(1+ct) \quad \text{if } a = 1 \\
\theta(t) &= \frac{p}{c(1-a)}[(1+ct)^{1-a} - 1] \quad \text{else}
\end{aligned}$$

This distribution encompasses the Poisson ($a = 0$), the Negative Binomial ($a = 1$) and the Poisson Inverse Gaussian ($a = \frac{1}{2}$) distributions.
Using infinite divisibility arguments (Maceda (1948), Feller (1971), see also Kestemont and Paris (1985) for a discussion), it can be shown that $N(t)$ may be interpreted as a Compound Poisson distribution. Therefore we can use Panjer's algorithm (Panjer (1981)) to write a recursion for the probabilities of $N(t)$:

$$(n+1)\Pi(n+1,t) = \frac{pt}{(1+ct)^a} \sum_{i=0}^{n} \frac{\Gamma(a+i)}{\Gamma(a)i!} \left(\frac{ct}{1+ct}\right)^i \Pi(n-i,t) \tag{5}$$

The mean and variance of $N(t)$ are given by

$$\begin{aligned} \mathrm{E}N(t) &= pt \\ \mathrm{V}arN(t) &= pt + pcat^2 \end{aligned}$$

The parameters are estimated by maximizing the loglikelihood :

$$l(p,c,a) = \sum_{i=0}^{\infty} N_i \ln\{\Pi(i,1)\}$$

where N_i is the number of policies with i claims.
Hürlimann (1990) shows that, for the Hofmann distribution among others, one has

$$E\hat{N}(1) = \hat{p} = \bar{N}$$

The other two parameters are found by a numerical maximization. We find

$$\begin{aligned} \hat{p} &= 0.15514 \\ \hat{c} &= 0.34853 \\ \hat{a} &= 0.44768 \\ l &= -54\,609.59 \end{aligned}$$

The fit is excellent as shows the following table

	Observed	Fitted
0	103704	103704,40
1	14075	14072,96
2	1766	1769,01
3	255	255,21
4	45	41,99
5	6	7,59
6	2	1,46

Table 2: Hofmann fit

The Chi-Square statistic is 0.4344 with 6 classes.

3. Non-parametric estimation

In the parametric case, we suppose that a U function has been choosen and that it only remains to estimate the parameters. Here we suppose a mixed Poisson distribution for which we don't specify a parametric distribution $U(\lambda)$ for Λ.
In that case, Simar (1976) shows that the maximum likelihood estimate of U will be attained for a discrete distribution function $U(\lambda)$ with a maximum number m of growing points.
The probabilities $\Pi(n,t)$ are then given by

$$\Pi(n,t) = \sum_{j=1}^{m} p_j e^{-\lambda_j t} \frac{(\lambda_j t)^n}{n!} \tag{6}$$

with $\sum_{j=1}^{m} p_j = 1$, $p_j \geq 0 \quad \forall j$ and m, the number of support points, i.e. the number of homogeneous classes of risks.
We will suppose in the sequel that $0 \leq \lambda_1 < \lambda_2 \ldots < \lambda_m$.
Simar (1976) gives an algorithm to find the non-parametric maximum likelihood estimators for an automobile portfolio. Unfortunately the loglikelihood is not concave everywhere and the algorithm does not converge certainly to the global maximum. In particular, Simar (1976) did not verify the fact that

$$\mathrm{E}\hat{N} = \bar{N} \tag{7}$$

in his numerical example (this property is valid for all non-parametric mixtures of the exponential family (Lindsay (1995))).
The maximum likelihood can also be found using a classical Newton-Raphson technique. The property (7) can be used to simplify the procedure by reducing the number of parameters to be estimated.
About the number of masses points m, Simar (1976) shows that the maximum likelihood estimator will be unique under the following conditions

$$\begin{aligned} m &\leq \min\left(q, \left\lfloor \frac{N+2}{2} \right\rfloor\right) \qquad \text{if } \lambda_1 = 0 \\ m &\leq \min\left(q, \left\lfloor \frac{N+1}{2} \right\rfloor\right) \qquad \text{if } \lambda_1 > 0 \end{aligned}$$

For our portfolio, we find the following maximum likelihood estimation :

$$\begin{array}{ll} \lambda_1 = 0.05461 & p_1 = 0.56189 \\ \lambda_2 = 0.24599 & p_2 = 0.41463 \\ \lambda_3 = 0.95618 & p_3 = 0.02348 \\ \multicolumn{2}{c}{l = -54\,609.456} \end{array}$$

Of course, the log-likelihood is higher than in any parametric case.

The fit is excellent as the following table shows :

	Observed	Fitted
0	103704	103703,83
1	14075	14075,57
2	1766	1765,36
3	255	255,77
4	45	43,63
5	6	7,50
6	2	1,16

Table 3: Non-parametric fit

The Chi Square statistic is 0.1245 with 6 classes.
The non-parametric case gives a physical interpretation of the heterogeneity of the portfolio : 56% of the risks follow a Poisson distribution with parameter $\lambda = 0.05461$, 41% of the risks follow a Poisson distribution with parameter $\lambda = 0.24599$ and 2% of the risks follow a Poisson distribution with parameter $\lambda = 0.95618$.

4. Optimal bonus-malus system

The bonus-malus system depends only on the number of accidents caused by the insured in the past. In our model, it is easy to see that it is sufficient to consider the total number of accidents without reference to the history of the accidents :

$$
\begin{aligned}
& dU(\lambda | N(t) - N(t-1) = k_t, \ldots, N(1) - N(0) = k_1) \\
= \; & \frac{\mathrm{P}[N(t) - N(t-1) = k_t, \ldots, N(1) - N(0) = k_1 | \lambda] dU(\lambda)}{\mathrm{P}[N(t) - N(t-1) = k_t, \ldots, N(1) - N(0) = k_1]} \\
= \; & \frac{\frac{e^{-\lambda t}\lambda^k}{\prod_{j=1}^t k_j!} dU(\lambda)}{\int_0^\infty \frac{e^{-\lambda t}\lambda^k}{\prod_{j=1}^t k_j!} dU(\lambda)}
\end{aligned}
$$

where $k = \sum_{j=1}^t k_j$.
The premium for the first year is an a priori premium because there is no information concerning the risk :

$$\mathrm{E}N(1) = \mathrm{E}\Lambda$$

For the t^{th} year, as the history of the accidents is unimportant, we take into account the information which consists in the number of accidents during the first t years and the premium is :

$$
\begin{aligned}
\mathrm{E}[N(t+1) - N(t) | N(t) = k] &= \mathrm{E}(\Lambda | N(t) = k) \\
&= \frac{k+1}{t} \frac{\Pi(k+1, t)}{\Pi(k, t)}
\end{aligned}
$$

This expression is general. It reduces to

$$\frac{p + kc}{1 + ct} = p\frac{1}{1+ct} + \frac{k}{t}\frac{ct}{1+ct}$$

in the binomial negative case and is by far more simple to use than the formulae derived by Tremblay (1992) for the particular case of the Poisson Inverse Gaussian distribution.
Assuming that the first premium paid is 100, we can construct a bonus-malus table depending on k and t with the formula :

$$\frac{100}{\mathrm{E}\Lambda}\frac{k+1}{t}\frac{\Pi(k+1,t)}{\Pi(k,t)} \tag{8}$$

The following good properties justify the "optimal bonus-malus" denomination :

1. The system is financially balanced each year :
$$\sum_{k=0}^{\infty}\Pi(k,t)\mathrm{E}(\Lambda|N(t)=k)=\mathrm{E}\Lambda \quad \forall t$$
2. The more you cause accidents, the more the premium :
$$\mathrm{E}[\Lambda|N(t)=k+1]>\mathrm{E}[\Lambda|N(t)=k] \quad \forall t\,,k$$
3. The premium always decreases when no more accidents are caused :
$$\frac{d}{dt}\mathrm{E}[\Lambda|N(t)=k]\leq 0 \quad \forall t\,,k$$

The properties 2 and 3 are easily shown by using the Cauchy-Schwartz inequality.

5. Bonus malus system for charged premiums

Following Lemaire (1985) and Tremblay (1992), we can construct a bonus-malus system for charged premiums using an exponential utility function

$$\mu(x)=\frac{1}{\gamma}(1-e^{-\gamma x}) \quad \gamma>0$$

with the principle of zero utility.
Using the formulae (3) and (4) the a priori premium becomes (Gerber (1979)) :

$$\begin{aligned} P &= \frac{1}{\gamma}\ln \mathrm{E}[e^{\gamma N(1)}] \\ &= \frac{1}{\gamma}\ln \mathrm{E}[e^{w\Lambda}] \quad \text{where } w=e^{\gamma}-1 \end{aligned}$$

The a posteriori premium is given in the same way than in the previous section :

$$\begin{aligned} P &= \frac{1}{\gamma}\ln \mathrm{E}[e^{w\Lambda}|N(t)=k] \\ &= \frac{1}{\gamma}\ln\left\{\frac{1}{\Pi(k,t)}\int_0^{\infty}e^{w\lambda}e^{-\lambda t}\frac{(\lambda t)^k}{k!}\right\} \\ &= \frac{1}{\gamma}\ln\left\{\left(\frac{t}{t-w}\right)^k\frac{\Pi(k,t-w)}{\Pi(k,t)}\right\} \end{aligned}$$

By normalizing such that the first premium is 100, the bonus-malus table is constructed with the formula

$$100\,\frac{\ln\left\{\left(\frac{t}{t-w}\right)^k\frac{\Pi(k,t-w)}{\Pi(k,w)}\right\}}{\ln\{\Pi(0,-w)\}} \quad \text{with } w=e^{\gamma}-1 \tag{9}$$

Once again, this formula is more general and more simple than in Tremblay (1992).

6. Numerical examples in the parametric case

t\k	0	1	2	3	4	5	6	7	8
1	87,47	162,05	278,98	424,26	582,15	744,85	909,45	1074,90	1240,79
2	78,92	138,18	228,80	341,68	465,47	593,78	724,02	855,14	986,73
3	72,59	121,75	195,41	287,24	388,63	494,30	601,87	710,35	819,32
4	67,65	109,66	171,52	248,62	334,22	423,85	515,35	607,78	700,70
5	63,66	100,33	153,53	219,78	293,66	371,35	450,87	531,31	612,27
6	60,34	92,87	139,47	197,41	262,27	330,71	400,95	472,12	543,81
7	57,52	86,76	128,15	179,54	237,24	298,34	361,18	424,95	489,24
8	55,09	81,64	118,84	164,94	216,82	271,93	328,75	386,48	444,73
9	52,96	77,27	111,02	152,77	199,85	249,99	301,79	354,50	407,73
10	51,07	73,50	104,36	142,46	185,51	231,47	279,03	327,50	376,50
20	39,48	52,10	68,49	88,35	110,94	135,42	161,14	187,63	214,62
50	27,13	32,59	39,18	46,87	55,55	65,05	75,21	85,86	96,86
100	20,14	22,95	26,17	29,83	33,88	38,31	43,05	48,07	53,32
500	9,90	10,47	11,08	11,74	12,43	13,16	13,93	14,73	15,58
1000	7,27	7,55	7,86	8,17	8,50	8,84	9,20	9,57	9,95
5000	3,54	3,60	3,66	3,72	3,78	3,84	3,90	3,97	4,03
25000	1,72	1,73	1,75	1,76	1,77	1,78	1,79	1,80	1,82

Table 4: Optimal bonus-malus table with the Hofmann fit

The table 4 compares favourably with Lemaire's (1985) and Tremblay's (1992).
Due to the choice of the form of $\theta'(t)$,we always have

$$\lim_{t\to\infty} E(\Lambda|N(t)=0)=0$$

and also

$$\lim_{t\to\infty} E(\Lambda|N(t)=k)=0 \quad \forall k$$

This fact corresponds to the well-known observation that most of the drivers are in the cheapest class when the frequency is low.
It may be inacceptable that a driver pays a premium equal to zero but in practice there is no problem because the convergence to 0 is attained by far after the mean driving time. Moreover the problem may be solved by adding a constant premium in our model :

$$\theta'(t)=\delta+\frac{p}{(1+ct)^a} \tag{10}$$

In this case, the premium has the following asymptotic behaviour :

$$\lim_{t\to\infty} E(\Lambda|N(t)=k)=\delta \quad \forall k$$

So the driver always pays a minimum premium δ.
Note that the following property seems to be verified :

$$\frac{d^2}{dt^2}\mathrm{E}(\Lambda|N(t)=k)\geq 0 \quad \forall k$$

Our experience shows it is true for typical automobile portfolios.

The bonus-malus tables for charged premiums also compares favourably with Lemaire's (1985) and Tremblay's (1992).

t\k	0	1	2	3	4	5	6	7	8
1	86,94	162,63	281,80	429,78	590,36	755,66	922,83	1090,82	1259,24
2	78,16	137,83	229,38	343,39	468,26	597,60	728,81	860,90	993,42
3	71,72	120,97	194,96	287,18	388,93	494,89	602,70	711,41	820,59
4	66,73	108,66	170,54	247,67	333,25	422,79	514,17	606,45	699,22
5	62,72	99,22	152,28	218,36	292,01	369,41	448,61	528,71	609,31
6	59,39	91,71	138,07	195,73	260,25	328,30	398,11	468,83	540,04
7	56,58	85,57	126,68	177,73	235,02	295,66	358,02	421,28	485,04
8	54,15	80,44	117,33	163,06	214,51	269,13	325,42	382,62	440,32
9	52,03	76,08	109,51	150,86	197,49	247,14	298,40	350,56	403,23
10	50,16	72,32	102,85	140,56	183,15	228,60	275,64	323,56	371,99
20	38,69	51,10	67,23	86,77	109,00	133,10	158,40	184,46	211,01
50	26,55	31,90	38,35	45,89	54,40	63,72	73,67	84,11	94,90
100	19,69	22,44	25,60	29,17	33,14	37,47	42,12	47,03	52,17
500	9,67	10,23	10,83	11,47	12,15	12,86	13,61	14,40	15,22
1000	7,10	7,38	7,68	7,99	8,31	8,64	8,99	9,35	9,73
5000	3,46	3,51	3,57	3,63	3,69	3,75	3,81	3,88	3,94
25000	1,68	1,69	1,71	1,72	1,73	1,74	1,75	1,76	1,78

Table 5: Bonus-malus table for charged premiums with an Hofmann fit ; $\gamma = 0.25$

As the following table shows, even for unreasonable values of γ, the difference with table 4 is small.

t\k	0	1	2	3	4	5	6	7	8
1	82,46	164,77	297,94	462,39	639,07	819,90	1002,29	1185,34	1368,73
2	72,23	133,25	228,79	347,49	476,56	609,59	744,22	879,56	1015,26
3	65,26	113,82	187,97	280,33	381,66	486,74	593,42	700,84	808,65
4	60,09	100,45	160,83	236,09	319,23	405,91	494,18	583,20	672,63
5	56,06	90,60	141,39	204,68	274,98	348,62	423,82	499,79	576,17
6	52,79	82,98	126,73	181,19	241,95	305,86	371,30	437,52	504,15
7	50,07	76,88	115,25	162,94	216,34	272,72	330,59	389,23	448,31
8	47,75	71,88	106,00	148,34	195,90	246,27	298,10	350,70	403,73
9	45,75	67,68	98,37	136,39	179,20	224,68	271,57	319,23	367,32
10	44,00	64,10	91,96	126,42	165,29	206,70	249,49	293,04	337,02
20	33,54	44,49	58,77	76,09	95,80	117,14	139,53	162,58	186,04
50	22,82	27,45	33,05	39,59	46,97	55,06	63,70	72,75	82,11
100	16,88	19,24	21,96	25,04	28,45	32,18	36,19	40,42	44,84
500	8,27	8,75	9,26	9,81	10,39	11,00	11,64	12,31	13,02
1000	6,07	6,31	6,56	6,82	7,10	7,39	7,68	7,99	8,31
5000	2,95	3,00	3,05	3,10	3,15	3,21	3,26	3,31	3,37
25000	1,44	1,45	1,46	1,47	1,48	1,49	1,50	1,51	1,52

Table 6: Bonus-malus table for charged premiums with an Hofmann fit ; $\gamma = 1$

7. Numerical examples in the non-parametric case

In this case, the formulae (8) and (9) may be rewritten as

$$(11) \qquad (8) = \frac{100}{\mathrm{E}\Lambda} \frac{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{k+1}}{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{k}}$$

$$(12) \qquad (9) = \frac{100}{\ln\{\Pi(0,-w)\}} \frac{\sum_{j=1}^{m} p_j e^{-\lambda_j (t-w)} \lambda_j^{k}}{\sum_{j=1}^{m} p_j e^{-\lambda_j t} \lambda_j^{k}}$$

The asymptotic behaviour of the bonus-malus tables is described as following :
If $\lambda_1 > 0$:

$$\lim_{t\to\infty} (11) = \min_{\lambda_j} \lambda_j \frac{100}{\mathrm{E}\Lambda}$$

$$\lim_{t\to\infty} (12) = \min_{\lambda_j} \lambda_j \frac{w}{\ln \Pi(0,-w)}$$

If $\lambda_1 = 0$:

$$\begin{aligned}\lim_{t\to\infty}(11) &= \min_{\lambda_j>0}\lambda_j\frac{100}{\mathrm{E}\Lambda} \quad \text{if } k>0\\ &= 0 \quad \text{if } k=0\\ \lim_{t\to\infty}(12) &= \min_{\lambda_j>0}\lambda_j\frac{w}{\ln\Pi(0,-w)} \quad \text{if } k>0\\ &= 0 \quad \text{if } k=0\end{aligned}$$

t\k	0	1	2	3	4	5	6	7	8
1	87,49	161,69	280,16	439,86	553,82	598,51	611,62	615,11	616,02
2	78,96	138,44	221,85	358,64	504,79	581,46	606,84	613,86	615,70
3	72,49	124,00	184,71	283,52	434,88	550,55	597,42	611,32	615,04
4	67,19	114,03	162,88	227,86	354,19	499,79	579,42	606,24	613,70
5	62,64	106,19	150,05	192,90	280,92	428,59	547,01	596,25	611,00
6	58,68	99,34	141,84	172,93	227,49	348,05	494,39	577,24	605,60
7	55,22	92,99	135,77	161,88	194,39	276,37	421,81	543,26	595,01
8	52,20	86,97	130,54	155,54	175,84	224,97	341,40	488,73	574,94
9	49,58	81,23	125,50	151,50	165,91	193,52	271,28	414,83	539,33
10	47,32	75,82	120,36	148,46	160,62	176,09	221,83	334,65	482,90
11	45,38	70,77	114,99	145,74	157,66	166,92	191,92	266,14	407,76
12	43,73	66,12	109,39	142,99	155,80	162,18	175,48	218,60	327,93
13	42,33	61,90	103,61	140,01	154,43	159,71	166,90	190,15	261,10
14	41,15	58,11	97,72	136,70	153,19	158,36	162,53	174,62	215,44
15	40,15	54,75	91,82	133,01	151,93	157,53	160,31	166,55	188,40
16	39,32	51,81	86,02	128,89	150,52	156,93	159,17	162,47	173,72
17	38,62	49,24	80,41	124,36	148,91	156,40	158,56	160,43	166,13
18	38,04	47,03	75,08	119,43	147,04	155,87	158,19	159,40	162,30
19	37,56	45,14	70,10	114,14	144,87	155,27	157,93	158,88	160,38
20	37,15	43,52	65,52	108,57	142,36	154,57	157,71	158,60	159,43
50	35,21	35,23	35,33	35,78	37,77	45,96	72,33	116,59	145,89
100	35,20	35,20	35,20	35,20	35,20	35,20	35,21	35,22	35,28
500	35,20	35,20	35,20	35,20	35,20	35,20	35,20	35,20	35,20
1000	35,20	35,20	35,20	35,20	35,20	35,20	35,20	35,20	35,20

Table 7: Optimal bonus-malus table with non-parametric fit

The bonus-malus table with the non-parametric fit greatly differs from the parametric case tables.

In fact the form of the bonus-malus table reflects the discontinuity of Λ. The non-parametric fit shows that there are three classes of risks : those with $\lambda = 0,05461; \lambda = 0,24599$ and $\lambda = 0,95618$. We find those three classes in the bonus-malus table ; locally, the table has the same comportment than a table constructed with a simple Poisson process : the premiums are almost indistinguishable because of the lack of heterogeneity. We have

$$\begin{aligned}\frac{100}{\mathrm{E}\Lambda}\lambda_3 &= 616,33\\ \frac{100}{\mathrm{E}\Lambda}\lambda_2 &= 158,56\\ \frac{100}{\mathrm{E}\Lambda}\lambda_1 &= 35,20\end{aligned}$$

The somewhat curious comportment of our table and the three classes are very visible for very bad risks. The following graph shows the evolution of the premium for $k = 15$.

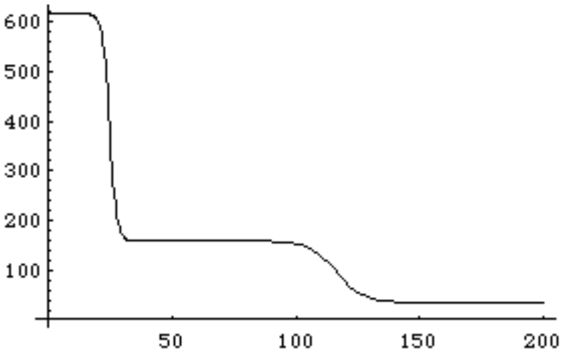

Figure 1: Optimal bonus-malus premium with non-parametric fit ; $k = 15$

Of course, because of that, the concavity of the premiums changes.
Such a curious comportment of the bonus-malus table seems very difficult to apply and so the non-parametric fit should not be used for the construction of bonus-malus tables.
The same comments apply for the charged premium bonus-malus table.

t\k	0	1	2	3	4	5	6	7	8
1	86,96	162,36	283,71	440,49	546,34	586,53	598,18	601,28	602,08
2	78,21	137,91	223,65	361,87	501,38	571,28	593,94	600,16	601,80
3	71,67	122,87	184,66	286,59	435,68	543,36	585,55	597,91	601,21
4	66,35	112,67	161,59	229,18	357,35	496,74	569,44	593,41	600,02
5	61,82	104,79	148,09	192,38	283,75	429,65	540,14	584,51	597,63
6	57,87	98,00	139,60	171,16	228,52	351,23	491,71	567,49	592,84
7	54,42	91,74	133,47	159,40	193,61	279,03	423,15	536,72	583,41
8	51,41	85,81	128,29	152,72	173,83	225,78	344,62	486,45	565,42
9	48,79	80,17	123,36	148,55	163,20	192,56	273,80	416,44	533,14
10	46,54	74,83	118,35	145,49	157,54	173,94	222,45	337,89	481,00
11	44,60	69,85	113,14	142,81	154,41	164,09	190,82	268,51	409,62
12	42,95	65,24	107,70	140,14	152,49	159,00	173,23	219,05	331,18
13	41,55	61,06	102,07	137,27	151,10	156,36	164,00	188,92	263,31
14	40,37	57,29	96,33	134,08	149,89	154,92	159,29	172,29	215,72
15	39,37	53,95	90,56	130,53	148,66	154,06	156,91	163,61	187,04
16	38,54	51,02	84,87	126,56	147,31	153,45	155,68	159,21	171,32
17	37,84	48,46	79,36	122,19	145,77	152,93	155,03	157,00	163,15
18	37,25	46,25	74,10	117,43	143,99	152,41	154,65	155,89	159,01
19	36,77	44,36	69,18	112,31	141,91	151,83	154,38	155,33	156,94
20	36,36	42,75	64,64	106,90	139,50	151,17	154,17	155,03	155,92
50	34,41	34,43	34,53	34,98	36,98	45,19	71,38	114,68	142,88
100	34,40	34,40	34,40	34,40	34,40	34,40	34,41	34,42	34,48
500	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40
1000	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40
5000	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40	34,40

Table 8: Bonus-malus table for charged premiums with non-parametric fit ; $\gamma = 0.25$

8. The stationary and transient distributions of the policyholders in a bonus-malus system with finite number of classes

We have seen in the previous sections the advantadge of using a parametric Mixed Poisson model to construct optimal bonus-malus systems. The non-parametric model was judged too discontinuous to give a nice form for the premiums. However this kind of adjustment for Mixed Poisson distributions presents high interest when one has to evaluate the mean of difficult functions over the risk's portfolio.
Let us assume $f(\lambda)$, a complicated function of λ. If we are interested in

$$\mathrm{E}f(\Lambda) = \int_0^\infty f(\lambda)dU(\lambda) \tag{13}$$

even a numerical integration may be untractable.
However, it is clear that using the non-parametric structure function of Λ will be more efficient. We know it gives a better fit and there won't be numerical integration needed. Only a convex combination of some $f(\lambda)$ will be performed. Equation (13) becomes

$$\mathrm{E}f(\Lambda) = \sum_{j=1}^{m} f(\lambda_j)p_j \tag{14}$$

which most of the time will be a sum with 3 or 4 terms.
Let us apply this technique to find the stationary and transient distributions of a typical Markov chain used in automobile insurance.
The bonus-malus systems with finite number of classes are used for a long time in most European countries. These bonus-malus systems are characterized by $s+1$ classes with growing premium percentages C_i, $i = 0, \ldots, s$. The movements of the drivers between the classes is given by transition rules depending on the number of accidents caused during one year. Most of the time, these transition rules give the model the Markov property. Even when it is not the case, a redefinition of the classes can give the Markov property to the model (see Lemaire (1985) for an application to the Belgian bonus-malus system).
For illustration, we will use the following bonus-malus system :

s	0	1	2	3	4	5	6	7	8
C_s	75	80	90	95	100	150	170	185	250

Table 9: Percentage premiums

Transition rules : -1 if no accident during the year
+3 per accident during the year

$$Y_t = \begin{cases} 3(N(t) - N(t-1)) & \text{if} \quad N(t) - N(t-1) = 1, 2, \ldots \\ -1 & \text{if} \quad N(t) - N(t-1) = 0 \end{cases}$$

Y_t are random variables representing the transition rules. Under the hypothesis that the Y_t are independent random variables, Dufresne (1988) has given a recursive technique to find the stationary distribution of the risks in the bonus-malus system.

The independence condition means that the general Mixed Poisson process can not be used for $N(t)$. However a Poisson process is adequate. Dufresne (1995) has used the Poisson distribution to find recursively, as a function of λ, the stationary distribution of risks with mean λ. He finds extremely complicated functions of λ and a software handling symbolic computations is welcome.
Then with this expression of the stationary distribution, $F^{\lambda}(x,s)$ where x is the class, Dufresne (1995) proposes to find the unconditional stationary distribution by

$$F(x,s) = \int_0^{\infty} F^{\lambda}(x,s)dU(\lambda)$$

As we mentioned above, this integration will be very easy if one has a non-parametric fit for $N(t)$. Moreover, as we know that only a few values of λ are needed, the algorithm of Dufresne (1988) may be forgotten at the benefit of the traditional technique of norming the left eigenvector of the transition probability matrix. This is the method we finally use to calculate the stationary distribution of the policyholders in our bonus-malus system. Note that all this remains true for the transient distribution functions.

For our bonus-malus system, the transition probability matrix is given by :

	0	1	2	3	4	5	6	7	8
0	$p(0)$	0	0	$p(1)$	0	0	$p(2)$	0	$p(\geq 3)$
1	$p(0)$	0	0	0	$p(1)$	0	0	$p(2)$	$p(\geq 3)$
2	0	$p(0)$	0	0	0	$p(1)$	0	0	$p(\geq 2)$
3	0	0	$p(0)$	0	0	0	$p(1)$	0	$p(\geq 2)$
4	0	0	0	$p(0)$	0	0	0	$p(1)$	$p(\geq 2)$
5	0	0	0	0	$p(0)$	0	0	0	$p(\geq 1)$
6	0	0	0	0	0	$p(0)$	0	0	$p(\geq 1)$
7	0	0	0	0	0	0	$p(0)$	0	$p(\geq 1)$
8	0	0	0	0	0	0	0	$p(0)$	$p(\geq 1)$

Table 10: Transition probability matrix

where $p(k) = \mathbf{P}[N(t+1) - N(t) = k]$ with $N(t) \sim Po(\lambda t)$.
From section 3, we know that

- $56,187\%$ of the policyholders have a frequency $\lambda = 0,05461$
- $41,464\%$ of the policyholders have a frequency $\lambda = 0,24600$
- $2,348\%$ of the policyholders have a frequency $\lambda = 0,95619$

Therefore we can calculate the stationary probability vector for those values of λ as well as their weighted mean which is the stationary probability vector of the portfolio. We find :

	$\lambda = 0,05461$	$\lambda = 0,24600$	$\lambda = 0,95619$	Portfolio
0	$0,8278$	$0,2598$	$0,0005$	$0,5728$
1	$0,0464$	$0,0724$	$0,0008$	$0,0561$
2	$0,0490$	$0,0926$	$0,0022$	$0,0660$
3	$0,0518$	$0,1185$	$0,0057$	$0,0783$
4	$0,0095$	$0,0876$	$0,0145$	$0,0420$
5	$0,0075$	$0,0942$	$0,0369$	$0,0441$
6	$0,0052$	$0,0977$	$0,0939$	$0,0457$
7	$0,0014$	$0,0880$	$0,2386$	$0,0429$
8	$0,0009$	$0,0888$	$0,6066$	$0,0516$

Table 11: Transition probability vector

This technique solves a problem encountered by Coene and Doray (1996) who used simulation to find the stationary distribution of a portfolio with $N(t)$ negative binomial distributed.
By simply taking powers of the transition probability matrix, we also find transient probabilities.
If we suppose that all the policyholders are in the beginning class

$$x_0 = (0,0,0,0,1,0,0,0,0)$$

then their distribution after 10 years in given by :

	$\lambda = 0,05461$	$\lambda = 0,24600$	$\lambda = 0,95619$	Portfolio
0	$0,8042$	$0,2448$	$0,0008$	$0,5533$
1	$0,0379$	$0,0441$	$0,0008$	$0,0396$
2	$0,0797$	$0,1739$	$0,0035$	$0,1170$
3	$0,0493$	$0,0930$	$0,0057$	$0,0664$
4	$0,0081$	$0,0587$	$0,0140$	$0,0292$
5	$0,0078$	$0,0829$	$0,0368$	$0,0397$
6	$0,0099$	$0,1429$	$0,0959$	$0,0671$
7	$0,0014$	$0,0725$	$0,2369$	$0,0364$
8	$0,0012$	$0,0868$	$0,6053$	$0,0509$

Table 12: Transient (10 years) probability vector

9. Conclusion

In this paper we have clarified how to construct a bonus-malus table using the net premium principle or the principle of zero utility when working with mixed Poisson distributions. The formulae are easily derived using the Bayes theorem and the form of the mixed Poisson distribution.
The parametric mixed Poisson distribution we use is more general than the traditional Negative Binomial or the Poisson Inverse Gaussian distributions. It has the disadvantadge that

three parameters have to be estimated (however the experimental mean directly gives one parameter for the maximum likelihood estimation). The fit is slightly better than the NB or the PIG fits.
In terms of goodness of fit, the non-parametric case is better than our three parameters parametric distribution. The estimation of the parameters is not too difficult due to the simple form of the non-parametric mixed Poisson distribution.
About the construction of a bonus-malus table, the parametric approach should be preferred because of its "continuity". Indeed the continuous form of the distribution of Λ gives a nice form to the bonus-malus table while the discontinuity of the non-parametric distribution gives the bonus-malus table a curious comportment with local almost constant premiums.
The tables constructed with the parametric distribution may be used as a starting point for methods like the one proposed by Coene and Doray (1996) to "fit" bonus-malus tables with classes.
The non-parametric fit shows its interest when one needs to evaluate the mean of difficult functions of λ over the portfolio. It replaces a complicated numerical integration by a short summation. The calculation of the stationary and transient distributions of a portfolio for a bonus-malus system with finite number of classes is a typical example of the advantadge of the non-parametric fit. Other quantities like the mean asymptotic efficiency of Loimaranta (1972) can also be easily evaluated with the help of the non-parametric fit.

References

Bühlmann, H. (1970). *Mathematical Models in Risk Theory.* Springer-Verlag, Berlin.

Coene, G. and Doray, L.G. (1996). A Financially Balanced Bonus-Malus System. *Astin Bulletin*, 26:107–115.

Dufresne, F. (1988). Distributions Stationnaires d'un Système Bonus-Malus et Probabilité de Ruine. *Astin Bulletin*, 18:31–46.

Dufresne, F. (1995). The Efficiency of the Swiss Bonus-Malus System. *Bulletin of the Swiss Actuaries*, 1995(1):29–41.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications Vol II (3ed).* Wiley, New-York.

Gerber, H. (1979). *An Introduction to Mathematical Risk Theory.* University of Pennsylvania.

Hofmann, M. (1955). Uber zusammengesetzte poisson-prozesse und ihre Anwendungen in der Unfall versicherung. *Bulletin of the Swiss Actuaries*, 55:499–575.

Hürlimann, W. (1990). On Maximum Likelihood Estimation for Count Data Models. *Insurance : Mathematics and Economics*, 9:39–49.

Kestemont, R.M. and Paris, J. (1985). Sur l'Ajustement du Nombre de Sinistres. *Bulletin of the Swiss Actuaries*, 85:157–163.

Lemaire, J. (1985). *Automobile Insurance : Actuarial Models.* Kluwer-Nijhoff, Netherlands.

Lindsay, B. (1995). *Mixture Models : Theory, Geometry and Applications.* Pennsylvania State University.

Loimaranta, K. (1972). Some Asymptotic Properties of Bonus Systems. *Astin Bulletin*, 6:233–245.

Maceda, E. (1948). On the Compound and Generalized Poisson Distributions. *Annals of Mathematical Statistics*, 19:414–416.

Panjer, H.H. (1981). Recursive Evaluation of a Family of Compound Distributions. *Astin Bulletin*, 12:22–26.

Simar, L. (1976). Maximum Likelihood Estimation of a Compound Poisson Process. *The Annals of Statistics*, 4:1200–1209.

Tremblay, L. (1992). Using the Poisson Inverse Gaussian Distribution in Bonus-Malus Systems. *Astin Bulletin*, 22:97–106.

Acknowledgement

Research support from the contract "Projet d'Actions de Recherche Concertées" no. 93/98-164 is gratefully acknowledged.

Jean-François Walhin*•
José Paris*

- * Institut de Statistique
  Voie du Roman Pays, 20
  B-1348 Louvain-la-Neuve
  Belgium

- • Le Mans Assurances, Belgique
  Avenue Louise, 222
  B-1050 Bruxelles
  Belgium