USING MIXED POISSON PROCESSES IN CONNECTION WITH BONUS-MALUS SYSTEMS [1]

BY

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 literature. We also compare this parametric method with a non-parametric one that has not yet been used in the actuarial literature 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 processes, 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 processes. 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.

---

[1] This paper has been presented at the XXVIIIth ASTIN Colloquium, Cairns 10-13 August 1997

If we define $N(t)$ as the number of claims in $(0, t]$ with $\Pi(n, t) = P[N(t) = n]$ =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.

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, in Tremblay's setting, seems to be very complicated needing for example the use of modified Bessel functions which is in fact unnecessary.

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:

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

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

where $\Psi_X(s) = E[e^{sX}]$ 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).

TABLE 1

REFERENCE PORTFOLIO

	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

As the claim number $N(t)$ is a pure birth process, the estimation of the intensity of the process

$$E[N(t+1) - N(t)|N(t)]$$

needs data reported on a long period of time, a situation which is not frequently met in practice.

Due to the fact that we have at disposal data reported on one period of time, the model allows the estimation of the intensity and the construction of a tarification based only on the total number of claims reported to the company. This system is comparable to a bonus-malus system but with an infinite number of classes. Due to the stationarity of the process we have

$$EE[N(t+1) - N(t)|N(t)] = EN(1)$$

and the system is at the equilibrium which is not the case with the bonus-malus systems met in practice.

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

$$\Pi(0, t) = e^{-\theta(t)}$$

$$\Pi(n, t) = (-1)^n \frac{t^n}{n!} \Pi^{(n)}(0, t) \quad n = 1, 2, \ldots$$

$$\theta'(t) = \frac{p}{(1 + ct)^a} \quad p > 0, \ c > 0, \ a \geq 0$$

$$\theta(0) = 0$$

By integration, we have

$$\theta(t) = pt \qquad \text{if } a = 0$$

$$\theta(t) = \frac{p}{c} \ln(1 + ct) \qquad \text{if } a = 1$$

$$\theta(t) = \frac{p}{c(1-a)} \left[(1 + ct)^{1-a} - 1 \right] \qquad \text{else}$$

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 $N(t)$:

$$\Pi(0, t) = e^{-\theta(t)}$$

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

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

$$EN(t) = pt$$
$$VarN(t) = pt + pcat^2$$

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

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

where $\bar{N}$ is the experimental mean.

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

$$\hat{p} = 0.15514$$
$$\hat{c} = 0.34853$$
$$\hat{a} = 0.44768$$
$$l = -54\ 609.59$$

The fit is excellent as shows the following table

TABLE 2

HOFMANN FIT

	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

The Chi-Square statistic is 0.4344 with 6 classes.

The Hofmann's distribution is based on a model that is verified on the data we are studying (Poisson distribution rejected, infinite divisibility, Shaked theorem). So, by its own nature, it is not surprising that it leads to a good fit.

3. NON-PARAMETRIC ESTIMATION

In the parametric case, we suppose that a U function has been chosen 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 $\sum_{j=1}^{m} p_j = 1$, $p_j \geq 0$ $\forall j$ and m, the number of support points, i.e. the number of homogeneous classes of risks.

We will suppose in the sequel that $0 \leq \lambda_1 < \lambda_2 ... < \lambda_m$.

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

$$\widehat{EN} = \bar{N} \tag{7}$$

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 mass points m, Simar (1976) shows that the maximum likelihood estimator will be unique under the following conditions

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

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

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

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

$$\lambda_1 = 0.05461 \qquad p_1 = 0.56189$$
$$\lambda_2 = 0.24599 \qquad p_2 = 0.41463$$
$$\lambda_3 = 0.95618 \qquad p_3 = 0.02348$$
$$l = -54\,609.456$$

Of course, the loglikelihood is higher than in any parametric case.

Note that the maximum likelihood method described in this section gives more information than the simple good guy bad guy model of Lemaire (1985). Indeed the procedure gives the number of mass points needed to have the highest likelihood while the good guy bad guy model imposes two mass points.

The fit is excellent as the following table shows:

TABLE 3

NON-PARAMETRIC FIT

	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

The Chi Square statistic is 0.1245 with 6 classes. The fit is excellent because we have 5 free parameters for 7 classes. Note that the distribution is not infinitely divisible in this case.

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:

$$dU(\lambda|N(t) - N(t-1) = k_t, \ ..., \ N(1) - N(0) = k_1)$$
$$= \frac{P[N(t) - N(t-1) = k_t, \ ..., \ N(1) - N(0) = k_1|\lambda]dU(\lambda)}{P[N(t) - N(t-1) = k_t, \ ..., \ N(1) - N(0) = k_1]}$$
$$= \frac{\frac{e^{-\lambda t}\lambda^k}{\prod_{j=1}^{t} k_j!} dU(\lambda)}{\int_0^\infty \frac{e^{-\lambda t}\lambda^k}{\prod_{j=1}^{t} k_j!} dU(\lambda)}$$

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:

$$EN(1) = E\Lambda$$

For the tth 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:

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

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:

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

The following good properties justify the 'optimal bonus-malus' denomination:
1. The system is financially balanced each year:

$$\sum_{k=0}^{\infty} \Pi(k, t)E(\Lambda|N(t) = k) = E\Lambda \qquad \forall t$$

2. The more you cause accidents, the more the premium:

$$\mathrm{E}[\Lambda | N(t) = k + 1] > \mathrm{E}[\Lambda | N(t) = k] \qquad \forall t, k$$

3. The premium always decreases when no more accidents are caused:

$$\frac{d}{dt} \mathrm{E}[\Lambda | N(t) = k] \le 0 \qquad \forall t, k$$

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

5. BONUS-MALUS SYSTEM FOR LOADED PREMIUMS

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

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

with the principle of zero utility.

Using the formulae (3) and (4) the a priori premium becomes (Gerber (1979)):

$$P = \frac{1}{\gamma} \ln \mathrm{E}\left[e^{\gamma N(1)}\right]$$

$$= \frac{1}{\gamma} \ln \mathrm{E}\left[e^{w\Lambda}\right] \qquad \text{where } w = e^{\gamma} - 1$$

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

$$P = \frac{1}{\gamma} \ln \mathrm{E}\left[e^{w\Lambda} | N(t) = k\right]$$

$$= \frac{1}{\gamma} \ln \left\{ \frac{1}{\Pi(k, t)} \int_{0}^{\infty} e^{w\lambda} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dU(\lambda) \right\}$$

$$= \frac{1}{\gamma} \ln \left\{ \left(\frac{t}{t - w}\right)^k \frac{\Pi(k, t - w)}{\Pi(k, t)} \right\}$$

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

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

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

6. NUMERICAL EXAMPLES IN THE PARAMETRIC CASE

With the expected value principle formulae, we find

TABLE 4

OPTIMAL BONUS-MALUS TABLE WITH THE HOFMANN FIT

$t \backslash k$	0	1	2	3	4	5
1	87,47	162,05	278,98	424,26	582,15	744,85
2	78,92	138,18	228,80	341,68	465,47	593,78
3	72,59	121,75	195,41	287,24	388,63	494,30
4	67,65	109,66	171,52	248,62	334,22	423,85
5	63,66	100,33	153,53	219,78	293,66	371,35
6	60,34	92,87	139,47	197,41	262,27	330,71
7	57,52	86,76	128,15	179,54	237,24	298,34
8	55,09	81,64	118,84	164,94	216,82	271,93
9	52,96	77,27	111,02	152,77	199,85	249,99
10	51,07	73,50	104,36	142,46	185,51	231,47
20	39,48	52,10	68,49	88,35	110,94	135,42
50	27,13	32,59	39,18	46,87	55,55	65,05
100	20,14	22,95	26,17	29,83	33,88	38,31

The table 4 is comparable 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

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

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 unacceptable 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:

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

So the driver always pays a minimum premium δ.

Note that the following property seems to be verified:

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

Our experience shows it is true for typical automobile portfolios.

The bonus-malus tables for loaded premiums, obtained with the zero-utility principle, are also comparable with Lemaire's (1985) and Tremblay's (1992).

TABLE 5

BONUS-MALUS TABLE FOR LOADED PREMIUMS WITH AN HOFMANN FIT; $\gamma = 0.25$

$t\backslash k$	0	1	2	3	4	5
1	86,94	162,63	281,80	429,78	590,36	755,66
2	78,16	137,83	229,38	343,39	468,26	597,60
3	71,72	120,97	194,96	287,18	388,93	494,89
4	66,73	108,66	170,54	247,67	333,25	422,79
5	62,72	99,22	152,28	218,36	292,01	369,41
6	59,39	91,71	138,07	195,73	260,25	328,30
7	56,58	85,57	126,68	177,73	235,02	295,66
8	54,15	80,44	117,33	163,06	214,51	269,13
9	52,03	76,08	109,51	150,86	197,49	247,14
10	50,16	72,32	102,85	140,56	183,15	228,60
20	38,69	51,10	67,23	86,77	109,00	133,10
50	26,55	31,90	38,35	45,89	54,40	63,72
100	19,69	22,44	25,60	29,17	33,14	37,47

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

TABLE 6

BONUS-MALUS TABLE FOR LOADED PREMIUMS WITH AN HOFMANN FIT; $\gamma = 1$

$t\backslash k$	0	1	2	3	4	5
1	82,46	164,77	297,94	462,39	639,07	819,90
2	72,23	133,25	228,79	347,49	476,56	609,59
3	65,26	113,82	187,97	280,33	381,66	486,74
4	60,09	100,45	160,83	236,09	319,23	405,91
5	56,06	90,60	141,39	204,68	274,98	348,62
6	52,79	82,98	126,73	181,19	241,95	305,86
7	50,07	76,88	115,25	162,94	216,34	272,72
8	47,75	71,88	106,00	148,34	195,90	246,27
9	45,75	67,68	98,37	136,39	179,20	224,68
10	44,00	64,10	91,96	126,42	165,29	206,70
20	33,54	44,49	58,77	76,09	95,80	117,14
50	22,82	27,45	33,05	39,59	46,97	55,06
100	16,88	19,24	21,96	25,04	28,45	32,18

7. NUMERICAL EXAMPLES IN THE NON-PARAMETRIC CASE

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

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

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

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

$$\lim_{t\to\infty}(11) = \min_{\lambda_j} \lambda_j \frac{100}{E\Lambda}$$
$$\lim_{t\to\infty}(12) = \min_{\lambda_j} \lambda_j \frac{w}{\ln\Pi(0,-w)}$$

If $\lambda_1 = 0$:

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

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

$$\frac{100}{E\Lambda}\lambda_3 = 616,33$$
$$\frac{100}{E\Lambda}\lambda_2 = 158,56$$
$$\frac{100}{E\Lambda}\lambda_1 = 35,20$$

J.F. WALHIN AND J. PARIS

TABLE 7

OPTIMAL BONUS-MALUS TABLE WITH NON-PARAMETRIC FIT

$t \backslash k$	0	1	2	3	4	5
1	87,49	161,69	280,16	439,86	553,82	598,51
2	78,96	138,44	221,85	358,64	504,79	581,46
3	72,49	124,00	184,71	283,52	434,88	550,55
4	67,19	114,03	162,88	227,86	354,19	499,79
5	62,64	106,19	150,05	192,90	280,92	428,59
6	58,68	99,34	141,84	172,93	227,49	348,05
7	55,22	92,99	135,77	161,88	194,39	276,36
8	52,20	86,97	130,54	155,54	175,84	224,97
9	49,58	81,23	125,50	151,50	165,91	193,52
10	47,32	75,82	120,36	148,46	160,62	176,09
11	45,38	70,77	114,99	145,74	157,66	166,92
12	43,73	66,12	109,39	142,99	155,80	162,18
13	42,33	61,90	103,61	140,01	154,43	159,71
14	41,15	58,11	97,72	136,70	153,19	158,36
15	40,15	54,75	91,82	133,01	151,93	157,53
16	39,32	51,81	86,02	128,89	150,52	156,93
17	38,62	49,24	80,41	124,36	148,91	156,40
18	38,04	47,03	75,08	119,43	147,04	155,87
19	37,56	45,14	70,10	114,14	144,87	155,27
20	37,15	43,52	65,52	108,57	142,36	154,57
50	35,21	35,23	35,33	35,78	37,77	45,96
100	35,20	35,20	35,20	35,20	35,20	35,20

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 $k = 15$.

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. The surprising results of the non-parametric method are due to the fact that the estimation of the distribution function U of the random variable Λ is based only on the observation of $N(1)$. Even if the period of observation is longer, the trouble will remain because as the frequency is low, the number of points on increase of U is always low and the number of classes of risks is low.
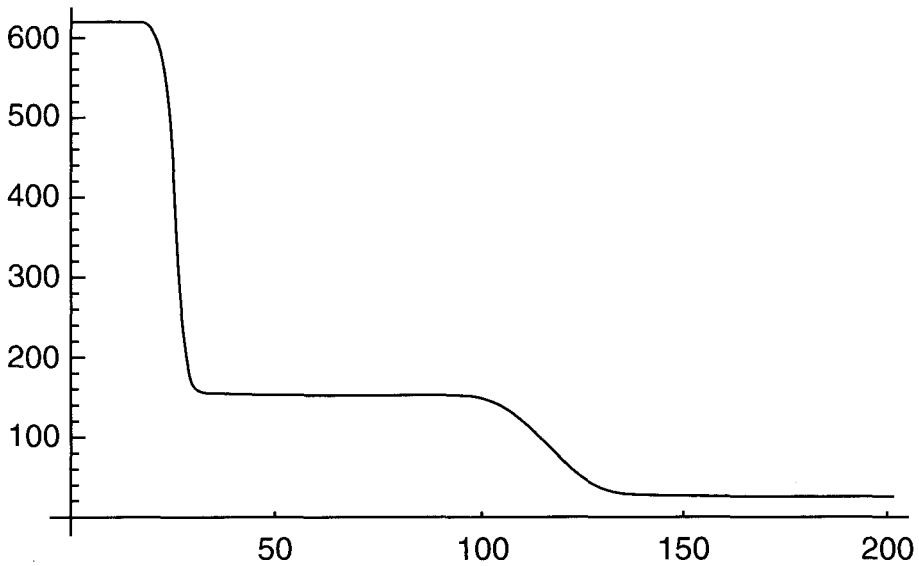
FIGURE 1: Optimal bonus-malus premium with non-parametric fit; $k = 15$

TABLE 8

BONUS-MALUS TABLE FOR LOADED PREMIUMS WITH NON-PARAMETRIC FIT; $\gamma = 0.25$

$t \backslash k$	0	1	2	3	4	5
1	86.96	162.36	283.71	440.49	546.34	586.53
2	78.21	137.91	223.65	361.87	501.38	571.28
3	71.67	122.87	184.66	286.59	435.68	543.36
4	66.35	112.67	161.59	229.18	357.35	496.74
5	61.82	104.79	148.09	192.38	283.75	429.65
6	57.87	98.00	139.60	171.16	228.52	351.23
7	54.42	91.74	133.47	159.40	193.61	279.03
8	51.41	85.81	128.29	152.72	173.83	225.78
9	48.79	80.17	123.36	148.55	163.20	192.56
10	46.54	74.83	118.35	145.49	157.54	173.94
11	44.60	69.85	113.14	142.81	154.41	164.09
12	42.95	65.24	107.70	140.14	152.49	159.00
13	41.55	61.06	102.07	137.27	151.10	156.36
14	40.37	57.29	96.33	134.08	149.89	154.92
15	39.37	53.95	90.56	130.53	148.66	154.06
16	38.54	51.02	84.87	126.56	147.31	153.45
17	37.84	48.46	79.36	122.19	145.77	152.93
18	37.25	46.25	74.10	117.43	143.99	152.41
19	36.77	44.36	69.18	112.31	141.91	151.83
20	36.36	42.75	64.64	106.90	139.50	151.17
50	34.41	34.43	34.53	34.98	36.98	45.19
100	34.40	34.40	34.40	34.40	34.40	34.40

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 advantage 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

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

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

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

which most of the time will be a sum with 3 of 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 numbers 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, $i = 0, ..., s$. 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:

TABLE 9

PERCENTAGE PREMIUMS

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

Transition rules: -1 if no accident during the year
$\qquad\qquad\qquad$ $+3$ per accident during the year
$\qquad\qquad\qquad$ $+0$ if any accident in class 8
$\qquad\qquad\qquad$ -0 if no accident in class 0

Let X_t be the class at time t. The transition rules Y_t is defined as

$$
Y_t = \begin{cases}
3(N(t) - N(t-1)) & N(t) - N(t-1) = 1,\ 2,\ \ldots \\
-1 & N(t) - N(t-1) = 0 \\
0 & N(t) - N(t-1) > 0 \quad X_{t-1} = 8 \\
0 & N(t) - N(t-1) = 0 \quad X_{t-1} = 0
\end{cases}
$$

We have $X_t = X_{t-1} + Y_t$.

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.

Note that the Dufresne's technique would not be applicable to bonus-malus systems with nonuniform penalties per claim while the technique described hereunder remains applicable for every 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) proposed 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:

TABLE 10

TRANSITION PROBABILITY MATRIX

	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	$p(0)$	$p(\geq 1)$

where $p(k) = P[N(t + 1) - N(t) = k]$ 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:

TABLE 11

STATIONARY PROBABILITY VECTOR

	$\lambda = 0,05461$	$\lambda = 0,24600$	$\lambda = 0,95619$	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

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 is given by:

TABLE 12

TRANSIENT (10 YEARS) PROBABILITY VECTOR

	$\lambda = 0,05461$	$\lambda = 0,24600$	$\lambda = 0,95619$	*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

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 disadvantage 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 advantage 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.

ACKNOWLEDGEMENT

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.

GELLER, 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: 157-164.

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.

JEAN-FRANÇOIS WALHIN
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

JOSÉ PARIS
Institut de Statistique
Voie du Roman Pays, 20
B-1348 Louvain-La-Neuve
Belgium